

The moment of truth for high-energy physics

Governments on both sides of the Atlantic seem under-prepared to defend the Large Hadron Collider from its critics in the US Congress.

The lengthy process whereby the United States Congress will, it is hoped, eventually provide funds for US participation in the Large Hadron Collider (LHC) at the European Laboratory for Particle Physics (CERN) has got off to an inauspicious start.

The trouble started in March when Joe Barton (Republican, Texas), in whose district sit the abandoned tunnels of the Superconducting Super Collider (SSC), pledged to do everything in his power, as he so vividly put it, to prevent "one thin dime from being spent on this".

It worsened two weeks ago, when the Science Committee of the House of Representatives — a forum normally sympathetic to both science and internationalism — voted to hold back \$35 million for the LHC pending clarifications from the US Department of Energy (DoE). James Sensenbrenner, chair of the Science Committee, has since visited CERN and returned with a lengthy wish list of modifications to the preliminary agreement which the DoE reached with CERN in February (see page 3).

Although European governments are not going to make a commitment to support some hypothetical, future US particle physics project, for example, it should not be beyond the wit of the DoE and CERN to satisfy Sensenbrenner. But the US administration must get squarely behind the project and sell it to the Congress.

In an influential report three years ago, Sidney Drell, deputy director of the Stanford Linear Accelerator Center, recommended that the United States should find resources from within its existing high-energy physics budget to participate in LHC. But he added an important proviso: without serious attention at the highest level — the White House — "international collaboration doesn't fly". Since then, the White House has been entirely silent on the issue. The administration has never bothered to explain to the American people that the

sum to be spent on this project will not go to Switzerland, as Barton asserts, but will be spent at Fermilab and Brookhaven, and with American industry. It has not countered the pernicious fairy-tale that lack of foreign support killed the SSC — the culprits were budgetary pressures, regional tensions and managerial incompetence.

In the previous Congress, officials hinted that silence was indeed their strategy, on the grounds that presidential support would single out the LHC for attack from the Republican Congress. But now that the project is under attack in the Congress, it is time for Jack Gibbons, the president's science adviser, to muster his boss in its defence.

Fortunately, Sensenbrenner and his Democrat counterpart George Brown, who shares some of his concerns, both agree that this project should proceed. They are aware of its importance to US particle physics, which is barely off its knees from the SSC debacle and would be humiliated and crippled by withdrawal from the LHC.

What Sensenbrenner and Brown are less sure about is the ability of the DoE to negotiate a good deal. Sensenbrenner says that he got more information in a day's visit to CERN than the department had provided him in the previous two years. Federico Peña, the new energy secretary, could demonstrate his competence to the Congress and his commitment to the science community by striking a revised deal acceptable to both CERN and to the Science Committee.

International collaboration in science is extraordinarily difficult to sustain. Nationalistic forces, be they angry Texan congressmen or hard-pressed Germans, will always attack budgets at their weakest point — wherever money seems to go abroad. US government officials who recognize the economic sense of sharing the costs of the LHC, as well as the physicists who anticipate its fruits, had better prepare for a lengthy battle in its defence. □

Telegrams abused

Circulars intended to notify astronomers of new detections should not be used to stake researchers' claims.

The International Astronomical Union (IAU) and the Smithsonian Astrophysical Observatory run an invaluable programme that rapidly distributes to astronomers around the world information on new and exciting astronomical sources. The intention is that the IAU circulars should be used to communicate positions and brightness of objects such as comets and supernovae, so that other astronomers can quickly confirm the data, or extend the observations to other wavelengths, in order to learn more about transient phenomena.

A flurry of circulars over the past six months or so threatens to overwhelm the administration of the office at Harvard University responsible for issuing them. Moreover, the purpose of the circulars is being subverted by some researchers seeking personal glory. These astronomers are submitting circulars that are miniature papers, containing interpretation and speculation, with the apparent intention of grabbing immediate publicity, without subjecting the work to the

peer-review process that has evolved as the scientific community's means of policing itself. Others are submitting circulars reporting data that cannot be obtained by anyone else and are irrelevant to observations at any other wavelength.

The IAU circulars serve astronomers excellently when they report data with the intention of promoting scientific understanding. Gamma-ray bursters are a good example. Rapid, accurate positions of the bursts are essential to finding a counterpart at optical or radio wavelengths, as demonstrated very recently. But vital information could be delayed — or even lost — amid a flood of inappropriate circulars speculating about the nature of a counterpart.

Circulars should have but one role: discoverers of unusual phenomena alerting others to their discoveries. Astronomers should submit circulars that do no more than help others to make observations, and the Central Bureau for Astronomical Telegrams at Harvard should reject circulars that do not serve this purpose. □



Key legislator collides with US officials on backing for CERN

[WASHINGTON] US plans to contribute half a billion dollars to the proposed Large Hadron Collider (LHC) at the European Laboratory for Particle Physics (CERN) in Switzerland have come under fire from the head of a key congressional committee, who claims that they take insufficient account of US interests.

Specific demands being made by James Sensenbrenner (Republican, Wisconsin), chair of the Science committee in the House of Representatives, include an assurance that US scientists will continue to enjoy free access to CERN, and a clear statement from European governments that, if a large accelerator is built in the future in the United States, they will help to pay for it.

He wants changes in a preliminary agreement reached earlier this year by CERN and the US Department of Energy (DoE), and says that US contributions will not be released until the agreement is modified.

Sensenbrenner has just returned from a private visit to CERN. Explaining his stance to a meeting of the American Association for the Advancement of Science (AAAS) in Washington DC last week, he said that he wants to see contingencies made for any cost

overruns in the construction of the LHC.

He also expressed concern about what he said was the lack of a formal role for the United States in managing the project, and about the impact of the proposed US contribution on existing high-energy physics facilities in the United States.

Sensenbrenner also presented his concerns directly to Federico Peña, the new energy secretary. Peña said after the meeting that the concerns were "reasonable", and that some of them were, "frankly, concerns of mine too".

According to Peña, the DoE is already discussing such issues with CERN. He said that this would "not necessarily" involve renegotiating an agreement in a way that would require endorsement by the full CERN council.

But issues such as the management of the LHC project and contingency funding may be difficult to address without such a renegotiation. Last month, Martha Krebs, the assistant energy secretary who negotiated the preliminary agreement for the US side before Peña's arrival at the department, said that although the deal could still be open to "technical amendments", this would



Sensenbrenner: Threatens to block funds until preliminary agreement is modified.

involve "nothing substantive".

Christopher Llewellyn Smith, the director-general of CERN, says Sensenbrenner was "doing his job" by scrutinizing the agreement closely. "He's asking good questions, but I think the answers are all there in the agreement," Llewellyn Smith says that an article in the agreement already guarantees open access to the LHC for US scientists. He said that costs are under firm control, and that the European members of CERN would be responsible for any overruns. "The US contribution is fixed," he says.

Llewellyn Smith adds that any assurances on support for future US projects could come only from European governments, not from CERN itself. He also says that any substantial modifications to the agreement would have to go back to the CERN council, which next meets in June, for ratification. "The deal can be changed, but we consider it to be satisfactory for all parties," he says.

As an authorization committee, the House Science committee has only limited power over budget matters. But, for political and technical reasons, it is highly unlikely that the appropriations subcommittees, which have the budgetary power, will fund a major new science project without the Science committee's support. This is especially true of LHC, which has hardly any constituency outside the science community.

Sensenbrenner's stance means that \$35 million that the Clinton administration had planned to spend on the LHC in the next financial year, beginning on 1 October, remains excluded from an advisory budget bill passed by his committee two weeks ago (see *Nature* 386, 748; 1997).

Officials at Fermilab and Brookhaven National Laboratory, where most of the money would be spent, are increasingly concerned that this money may remain excluded from the budget. They had hoped Sensenbrenner would return from CERN satisfied, and would amend the committee's proposal

House limits reform to patent laws

[WASHINGTON] Legislation designed to overhaul the US patent system, bringing it more into line with international practice, passed the House of Representatives last week, but only after last-minute amendments weakened several key reforms.

The bill, designated HR 400 and sponsored by Howard Coble (Republican, North Carolina), converts the Patent and Trademark Office from a federal agency into a government-owned corporation financed by patent fees. It also requires that patent applications be published 18 months after being filed, as is the practice in Europe and Japan. At present, US applications become public information only after a patent is granted, usually 20 to 22 months after filing.

The 18-month rule is intended to reduce the surprise element in the award of patents, and force the early publication of novel ideas. But small businesses, individual inventors and universities will be exempt, thanks to an amendment sponsored by Mary Kaptur (Democrat, Ohio).

Two other clauses had been of particular concern to universities. One would have allowed the challengers of a patent to participate fully in the 're-examination'

process. Critics claim that this would have favoured wealthy companies over small patent holders.

The second was a clause allowing those who had been making prior commercial use of an idea to continue using it even after a patent is granted to someone else. Some in the research community fear that this condition would make it more difficult for universities to grant exclusive licences for their inventions, which have become an important source of revenue.

The re-examination clause was cut out by an amendment approved during the debate on the floor of the House last week. The allowance for 'prior use' was restricted to satisfy the university community and other opponents of the original bill.

A disappointed Coble said these and other changes to his legislation have weakened it, but "the damage is not irreparable". A similar bill, S 507, which is sponsored by Orrin Hatch (Republican, Utah), is now moving through the Senate. Representatives of the university community say that they are planning to monitor its progress closely. **Tony Reichhardt**

See Commentary, page 13.

before it reached the floor of the House.

But Sensenbrenner now says its "up to secretary Peña" to meet his demands and get the money restored. The bill is expected to reach the House floor later this month. According to congressional staff and laboratory officials, supporters of the LHC are worried that an attempt to restore the funds on the floor of the House could well be defeated. "Everyone knows that if you take this to the floor, you lose," says one.

Sensenbrenner's staff say that he supports US participation in the LHC, but wants a better deal that will survive future attacks in Congress. "There are good deals and bad deals — Martha Krebs' deal is a bad one that has to be changed," Sensenbrenner says.

Nevertheless, his recent public statements have been short on praise for the LHC, and have reflected some of the views of Joe Barton (Republican, Texas) the project's fiercest congressional critic (see *Nature* 386, 97; 1997).

Sensenbrenner told last week's AAAS meeting that, while he was "all for" international collaboration in science, the Superconducting Super Collider (SSC) had failed because it had not attracted international contributions, particularly from Europe.

"Carlo Rubbia, the [then] director of CERN, waved his finger at the United States and actively lobbied against European support for SSC," Sensenbrenner argued. A European pledge to help with a future US machine would "put an end to Rubbia-ism in Geneva," he suggested.

Whatever happens next, Sensenbrenner's attack on the initial deal between the DoE and CERN marks the unravelling of the Clinton administration's strategy of pushing through, with minimal fuss, the largest-ever US investment in an overseas science project. Costs would involve \$450 million from the DoE and probably another \$80 million from the National Science Foundation.

Three years ago, a panel chaired by Sidney Drell, deputy director of the Stanford Linear Accelerator Center in California, suggested that the United States should participate in the project. But it warned that it would need the explicit endorsement of President Clinton to succeed (see *Nature* 369, 266; 1994).

Since then, however, Krebs' Office of Energy Research has pursued the agreement on its own, with little public backing from Hazel O'Leary, the former energy secretary, or other senior members of the administration. Last week, the new energy secretary sought to correct this, speaking out in favour of the project at the AAAS meeting.

"We are requesting funds to participate in the construction of the next major high energy physics research facility in the world — the LHC," Peña said. "Successful international collaboration is important to the future of these large scientific facilities; we must succeed on this one."

Colin Macilwain

Spaceflight monkey's death 'raises new safety issues'

[WASHINGTON] US space researchers have dropped plans to fly rhesus monkeys on the Russian Bion 12 spacecraft next year. Their decision follows the conclusions of a review panel that the animals face a previously unsuspected risk of death on their return to Earth.

The panel was set up after one of the two monkeys that were used on the 14-day Bion 11 mission last December died unexpectedly under anaesthesia during a routine biopsy conducted the day after landing (see *Nature* 385, 289; 1997). The mission was one of a series designed to test the effects of weightlessness and space radiation on living organisms.

The decision brings to an abrupt end a long-running controversy about whether the National Aeronautics and Space Administration (NASA) should fly monkeys on the Bion missions. The December mission, for which NASA had followed the advice of several scientific review panels, had encountered protests from animal rights groups. There have also been attempts to kill the programme in Congress. Indeed, Tim Roemer (Democrat, Indiana), a member of the House of Representatives Science committee, had begun circulating a "Dear Colleague" letter headlined "Get the monkey off the taxpayer's back!", the day before NASA's decision.

Alice Hellerstein, a spokeswoman for the American Physiological Society, which had backed the Bion experiments during congressional debates last year, says "[t]he timing is unfortunate, because it looks like they were pressured politically". However, she believes that, despite the possible embarrassment to NASA, the agency made a "tough but honest" decision based on new scientific evidence.

Ronald Merrell, chairman of Yale University's department of surgery and head of the panel that reviewed the Bion research protocols both before and after the flight, says that as well as being a genuine surprise, the result is of considerable scientific interest.

Russian researchers had previously conducted 10 successful biopsies on monkeys after long space-flights, following procedures used routinely in laboratories on Earth. But they had never anaesthetized an animal so soon after its return from space.

According to NASA's chief veterinarian, Joseph Bielitzki, the soonest it had been done previously was seven days after landing. But the Bion 11 experiment required that tissue be taken immediately from the monkeys in order to investigate the body's transition from weightlessness to gravity.

Merrell says that physiological changes that occur in space appear to have reduced the monkey's ability to tolerate the stress of anaesthesia. The review panel therefore concluded that continuing with the same experimental protocol for Bion 12 posed an additional risk to the test subjects.

The panel did not tell NASA to scrap the project, says Merrell. But, in the light of the new findings, the agency faced a choice between risking the death of the Bion 12 monkeys — which would have inflamed political opponents — or waiting several days after landing before taking biopsy tissue, which would reduce its value.

NASA headquarters chose to cut its losses after what Bielitzki describes as "a lot of discussion", reaching as high as Daniel Goldin, NASA's administrator. According to Bielitzki, there is "a significant amount of disappointment" among Bion 12 researchers.

The rhesus monkeys were originally to have flown on a US Spacelab mission last year, but were switched to Bion when that flight was cancelled. The experiment had been approved by four scientific review panels since 1988, including one convened by the American Institute of Biological Sciences. But the plan to anaesthetize the monkeys immediately after landing never emerged as a significant worry. "Hindsight is always 20-20," says Bielitzki.

Other US animal experiments will still fly on Bion 12, along with investigations sponsored by Russian and French researchers. And, ironically, the death of the Bion 11 monkey may leave an important scientific legacy. NASA may want to reconsider whether or not to operate immediately on astronauts who return from space in an emergency. "[Agency doctors] are probably thinking about this right now," says Merrell.

Tony Reichhardt



THE THREE WISE NASA OFFICIALS

CJD link prompts ban on brain tissue use

[TOKYO] The Japanese health ministry has belatedly decided to ban the import and use of dura mater grafts, almost ten years after similar action was taken by the US Food and Drug Administration (FDA). The decision highlights major deficiencies in the monitoring for safety of health-care products in Japan.

The grafts, freeze-dried tissue from the outer skin of the human brain that is used in many neurosurgical procedures, are derived from human cadavers and have been linked to cases of Creutzfeldt-Jakob disease (CJD). Last week, an expert committee approved an earlier announcement by the Japanese government recommending that such grafts should be replaced by autogenic grafts or artificial substitutes.

The government's decision follows the results of an epidemiological study that found that infected dura mater could have caused as many as 43 cases of CJD in Japan over the past ten years. The study was commissioned by the ministry, and directed by Sato Takeshi of the National Center of Neurology and Psychiatry.

Although Japanese officials were informed by researchers in 1987 about a report on iatrogenic transmission of CJD by dura mater by the US Centers for Disease Control and Prevention (CDC), they claimed that they became aware of the issue only last year.

Most commercially available dura mater products are fabricated by two German companies, Braun Melsungen and Biodynamics. After 1987, many countries encouraged autografting and the establishment of tissue banks. Only Japan continued to rely heavily on imported dura mater products. The practice of autografting is still rare in Japan, and the only tissue banks in existence are run by research units of university hospitals.

Apart from a single recent Japanese case, where the origin of the graft could not be established, all known cases of transmission of CJD by dura mater have been linked to grafts fabricated by Braun Melsungen before 1987, sold under the name Lyodura.

In contrast, at least three of the Japanese CJD cases related to dura mater appear to have been caused by surgical interventions carried out after an international product recall was issued by Braun Melsungen in March 1989. A representative of BBS, which distributes Braun Melsungen's products in Japan, has been reported as saying that the product recall was not fully applied in Japan because of a linguistic misunderstanding.

The FDA issued a product recall and an import ban on dura mater manufactured by Braun Melsungen as soon as the CDC results became known in 1987.

Japanese experts appear to have known of the relatively high number of CJD cases relat-

ed to dura mater grafts in their country for some time. But the Japanese ministry took action only after an expert committee of the World Health Organization (WHO) recommended in March that dura mater "should no longer be used" on human patients.

The risk of transmission of CJD through allogenic dura mater has been widely acknowledged since the first such case was reported in the United States in 1987. But the risk of transmission is thought to have been considerably reduced by strict donor screening and quality control protocols, as well as by chemical and physical treatment aimed at inactivating the CJD agent.

Fewer than 20 cases of CJD linked to dura mater allografts had been reported worldwide before the preliminary results of the Japanese study were released last year. By then, 28 Japanese cases had already been identified. European experts say they are "surprised" by the new Japanese data, but that similar high rates of transmission elsewhere are highly unlikely.

The recommendation that all existing cadaveric dura mater products be banned has now been endorsed by both WHO and the Japanese health ministry. But companies that produce the material are resisting the move, claiming that it has not been scientifically demonstrated as necessary.

A spokesperson of Biodynamics, the world's second-largest provider of dura mater, describes the recommendation as "scandalous", and says that the company has urged WHO to reveal the scientific evidence.

The validity of the recommended ban is being investigated by both the US CDC and

the FDA. Apparent problems in the communication of scientific information by WHO seem to have been reflected in the fact that an initial version of the press release on the expert meeting on bovine spongiform encephalopathy, CJD and medical products stated mistakenly that "CJD has been found to be transmissible by blood". The statement was later corrected.

Some critics have used the situation in Japan to support their claim that the ministry appears unable to take its own decisions about the safety of medical products, even where there is solid scientific evidence of a potential danger. "They needed an official announcement by WHO to legitimize their action," says Shohei Yonemoto, an independent science policy observer in Tokyo. Yonemoto argues that the ministry needs to improve its decision-making and risk assessment procedures.

There is also a growing feeling that, at a time when public debates on organ transplantation are heating up in Japan (see panel below), the dura mater case reveals severe shortcomings in the regulation and handling of products derived from human donors.

Under present Japanese legislation, most medical products derived from human donors remain unregulated. Until recently, even products treated as "medical devices" — such as dura mater — were not subjected to regular screening and quality control procedures after initial approval.

And Japan still lacks both a legal framework and an independent organization to guide the handling of human tissues, including tissues used for research. **Robert Trione**

Japan takes step towards organ transplants

[TOKYO] After decades of debate, a bill to provide a legal framework for organ transplants was finally approved last Thursday (24 April) by the lower house of Japan's parliament, the Diet, in what appears to have been its first-ever free vote. The bill, introduced three years ago by an all-party group, still has to be ratified in the upper house, where its fate is by no means certain.

The bill makes it explicitly legal to transplant organs from "brain death bodies". But, although purporting to regulate organ transplants in general, the proposed law covers only main organs, and omits coronary arteries,

bone marrow and other tissues — all of which are in far greater demand than hearts or livers.

The use of organs from brain-dead donors has been a controversial issue. Organ transplants are not technically illegal. But no heart transplant has been carried out since Japan's first such operation ended in scandal some 27 years ago.

As a result, many Japanese in need of transplants have been going overseas, and there is a growing population of patients with renal failure who cannot get transplants and have to rely on dialysis.

Cultural factors contribute

to the widespread opposition to 'brain death' as the legal definition of death. But some also blame the medical profession. In most countries, organ transplants are not regulated by law but by professional guidelines. It was only last year that the Japanese Society for Organ Transplants started drafting guidelines.

The bill does not cover organs used for research. But it was recently revealed that Japanese university researchers have for at least the past six months been importing livers from brain-dead US donors for use in safety evaluations of pharmaceuticals. **R. T.**

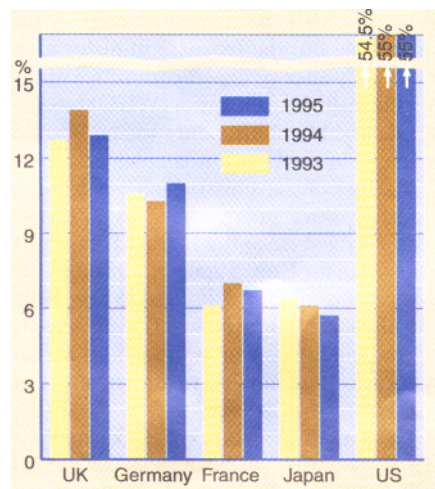
Structural biology gets UK funding boost

[LONDON] The UK government's Biotechnology and Biological Sciences Research Council (BBSRC) is to provide £8.5 million (US\$12.75 million) to help to set up multi-disciplinary centres and to update equipment in structural biology research, in which it says Britain is now second only to the United States.

The funds will be released over three years, and are being allocated in response to a review by the research council of structural biology research, published last week. This warns that the United Kingdom could lose its pre-eminent position in the discipline without greater collaboration between universities, and between universities and industry.

"Much of biology is changing; it is becoming a 'big science' which needs state-of-the-art equipment and can no longer be effectively carried out by small, isolated teams," says Chris Higgins, professor of clinical biochemistry at the University of Oxford, who chaired the BBSRC review. Other members of the panel included representatives from public and private funding agencies in the United Kingdom, the United States and Europe.

The multidisciplinary centres will be formed in partnership between the BBSRC, universities and the pharmaceutical industry, which is becoming increasingly dependent on research into the structure and function of biological molecules. Higgins says he hopes the extra funds can be matched by industry and other sources. A call for



Percentage shares of world output in structural biology, weighted for 'quality'.

proposals for the new research centres is being issued this week.

"We need a limited number of centres that provide both a focus for high quality research and access for other biology researchers. We want to avoid every university setting up structural biology departments," says Higgins. Colin Miles, head of the biomolecular sciences group at the BBSRC, says "the centres need not be bricks and mortar at a single location" — they could be "virtual centres".

The new funds will supplement the BBSRC's current £17–£20 million annual

spending on structural biology, 10 per cent of the council's total budget for biology research. A similar sum is spent by the Medical Research Council. Between them, the funding councils account for 80 per cent of UK spending on structural biology. The BBSRC has already spent £5 million on new equipment.

The council's report contains detailed bibliometric information on world structural biology research. Higgins describes it as "one of the most comprehensive and thorough surveys of any research discipline". It finds that the United Kingdom leads Europe in structural biology research, with the UK share of world output increasing steadily from 9 per cent in 1986 to 12 per cent in 1995.

Universities account for 70 per cent of research output, with research council institutes at just under 17 per cent and industry on 5.6 per cent. Structural biology research is concentrated at a handful of institutions. Just five — Oxford, Cambridge, Birkbeck (London), York and University College London — out of the United Kingdom's 140 universities account for 50 per cent of output.

The report points out that access to a synchrotron radiation facility is essential for structural biology. UK researchers now rely on the synchrotron radiation source at the Daresbury Laboratory in Warrington, Cheshire, a facility that will outlive its usefulness in three years. The BBSRC has not yet decided whether to find the £15 million needed to help finance a new source. **Ehsan Masood**

Korea leaps before it looks over gene therapy guidelines

[TOKYO] In an unusual move, a group of medical researchers in Seoul is expected to announce this week that they have carried out the first clinical application of gene therapy in South Korea. The announcement is surprising because South Korea, unlike most industrialized countries, has yet to establish guidelines for gene therapy.

Treatment began last month on a 33-year-old woman suffering from breast cancer at Samsung Medical Center, a lavish new hospital set up by the South Korean industrial conglomerate.

The programme involves a retroviral vector from the United States, and is a collaboration between researchers at the centre, led by Chan H. Park, a group at Seoul National University under Sunyoung Kim, and a group at the University of Pittsburgh Medical Center led by Michael Lotze.

Using a protocol originally developed by Lotze's group, skin fibroblasts from the patient were treated with a retrovirus containing the gene for interleukin-12, an agent believed to activate immune defences

against cancer cells. Cells containing the gene were then injected into the patient after irradiating them to stop them multiplying and overproducing interleukin-12, which can have toxic side-effects.

The proposal for the clinical trial went through three review steps over about 18 months. The protocol was first considered by an institutional review board at the Samsung Medical Center, then an *ad hoc* committee at the Ministry of Health and Social Welfare. Finally, the US Food and Drug Administration checked the protocol before approving the export of the gene-carrying retrovirus.

But unlike the United States and Japan, where the first applications of gene therapy went through exhaustive public debate and examination by committees of the ethical and social implications of gene therapy as well as its medical aspects, South Korea is still in the process of setting up guidelines for gene therapy. The issue has been publicly debated for several years, but a government committee to set up guidelines was

established only about two months ago, according to Kim, and it will be some time before they are approved.

Park says the government process was taking "too long". He argues that the protocol has already been extensively reviewed in the United States, and that its safety is therefore well established. He also says it would "not be fair to be left behind" because of the lack of government regulations, and argues that the trial has undergone extensive domestic as well as international review.

Both Kim and Park expect their announcement to be welcomed by the Korean public, a view echoed by other Korean medical researchers. Koreans are "less concerned" and "less aware" of any potential problems associated with gene therapy, says one professor at Seoul National University. But he admits to a concern that the announcement may generate strong competition among Korean clinicians "without consideration of ethical and scientific problems". **David Swinbanks**

Gene researchers promised relief from regulatory burden

[MUNICH] Germany's research minister, Jürgen Rüttgers, has promised to submit proposals to the cabinet suggesting how the conditions for genetic research could be improved, primarily by reducing administrative burdens on researchers and increasing the efficiency of funding.

Rüttgers' promise followed the publication last week of a report by the Research, Technology and Innovation Council, an *ad hoc* advisory group set up in 1995 to report to Chancellor Helmut Kohl on scientific issues.

After nearly a year's analysis, the council — which includes 26 representatives of the academic community and industry, as well as several politicians — reached familiar conclusions about the problems facing academic and industrial researchers using genetic engineering techniques in Germany.

Despite recent improvements (see *Nature* 379, 670; 1996), it found that researchers still face delays in gaining approval for experiments, there remain few sources of venture capital, and public hostility to genetic engineering has discouraged research at a time when it has been booming elsewhere.

The report makes 96 recommendations. These include concentrating public funds on a limited number of centres of excellence, and an increase in the number of foreign researchers in these centres. The report suggests more funding for bioinformatics, an area which it warns is dangerously weak in Germany.

The council also says that Germany should rethink what some consider to be an obsessively democratic attitude towards biotechnology, and remove the requirement for a public hearing for major industrial initiatives. Field trials of genetically manipulated crops should be extensively deregulated, it says. These have been a particular target for public hostility (see *Nature* 380, 94; 1996).

The council's report calls for tax breaks for industries conducting research, incentives for venture capitalists to invest in German biotechnology, a single — and cheap — European patent, and more training opportunities in genetic engineering for young scientists.

Detlev Ganten, director of Berlin's Max Delbrück Centre for Molecular Medicine and a member of the advisory council, says that, even if only some of the suggestions are implemented, Germany will have a better chance of catching up with its competitors. Rüttgers has promised to submit his proposals by the autumn.

The report is the second from the council. The first — on the information society — was published last year.

A.A.

Conservatives undermine Italy's university reforms

[MUNICH] Hopes for significant reforms to Italy's widely criticized system of academic promotion have been dashed — at least for the near future — by compromises made to satisfy the powerful conservative academic lobby within Italian politics.

Last week, the Italian Senate passed a university bill that fails to solve what many consider to be the central problem with the system — the lack of staff mobility between universities and the powerful influence that personal connections play in making academic appointments.

When Luigi Berlinguer, the minister for education and research and formerly a professor of law at the University of Siena, took office last year, one of his priorities was to change the system by which faculty members are appointed through centrally organized competitions, or *concorsi* (see *Nature* 382, 7; 1996).

Winners of these competitions are allocated to individual universities, and deals are often made to ensure that scientists are appointed to institutions with which they have strong connections.

Berlinguer and many others feel that this procedure undermines scientific competitiveness. His original bill proposed that non-Italians should take part in judging committees to reduce the influence of local connections. It said that the number of winners should exceed the places available, giving universities a wider selection of candidates. He also wanted universities to be barred from appointing a candidate from the same city to the post of professor.

But last week's Senate vote compromises considerably on each of these points. First, it no longer includes a requirement for foreign scientists to sit on judging committees.

Second, candidates would be able to apply for promotion, at any time, to a new three-member national committee, and repeatedly reapply if they do not succeed initially. Successful candidates would remain on a national list for eight years.



Berlinguer: Senate weakened plans.

Universities would be able to select faculty members from this list, but would not — at least initially — be required to recruit outsiders. A new clause introduced into the bill states that those added to the list within its first four years can be recruited locally for a further eight years.

"This means that local candidates, whatever their merit, will become professors, as the local selection committees are likely to choose a local candidate," says Piergiorgio Strata, a professor of neurophysiology at the University of Turin. Strata, an active campaigner for improvements to the Italian academic system, fears that a further generation of academics will be recruited in this way.

The draft law still has to be voted on by the Chamber of Deputies, Italy's lower parliamentary chamber, where it is likely to face a rough ride from conservatives. If it fails to win support, reformers are likely to face a long wait before the issue finds its way back to the top levels of the government's problem-filled agenda.

Alison Abbott

'Britain must tackle lack of trust on funding'

[LONDON] Britain's Royal Society has asked the winner of this week's general election to tackle the problem of the "increasing lack of trust" between researchers and funding agencies. It wants the new government to establish a "fresh balance" between the freedom of fundamental research teams to choose their own goals and the pressures of accountability in the use of public funds.

The demand is made in a document entitled 'Memorandum for an Incoming Government'. This argues that the perception among scientists of an increasing tendency for research programmes to be defined and directed "from the top down" has led to a situation in which university staff are "beginning to think of research councils as 'the enemy' rather than as agencies designed to help".

The statement makes no direct mention

of levels of funding for science. But it does warn whichever party forms the next government of the dangers arising from an increasingly short-term and 'dirigiste' attitude towards fundamental research. Both the ruling Conservative party and its Labour opponents, who were widely expected to win the election, remained discreetly silent about science funding during the election campaign (see *Nature* 386, 314; 1997).

The Royal Society asks for an end to "frequent, ideologically motivated reviews" of research establishments. It also offers the services of its 1,200 fellows as a source of scientific advice to the government, hoping that they could be used in a way analogous to the US government's use of the madvisory capacity of the National Academy of Sciences.

Help lies around the corner for Internet's frustrated users

Declan Butler

A flood of initiatives promises to relieve some of the current difficulties experienced with the Internet, particularly slow international links. In the process, faith in public systems of communication has been confirmed.

Faced with the growing frustration of users, governments and research organizations are moving to improve the Internet. One goal is to eliminate international bottlenecks; another is to develop the infrastructure and protocols needed to create a second-generation system.

Researchers can expect a significant acceleration in the speed of communication, the elimination of blockages and the development of new, high-performance applications. Any idea that such developments will emerge unprompted from market forces appears to have been abandoned.

In Europe, the improvements will be experienced soon. A new interconnect facility providing links among national research networks at a bandwidth of 34 megabits per second (Mbps) will be opened this month, known as the Ten-34 project (Trans-European Network Interconnect at 34 Mbps). In the United States, universities have recently launched a similar project — Internet-2 — to improve links between them.

Links have been neglected

The US National Science Foundation (NSF) is also funding a high-speed research backbone, the vBNS, which will link about 100 universities (see *Nature* 380, 93; 1996). In parallel, President Bill Clinton has pledged

\$100 million to the Next Generation Internet Initiative. This will create high-performance networks for federal agencies such as NASA and the National Institutes of Health.

Another event planned this month is a meeting of the G7 group of industrialized countries to discuss proposals for linking up national high-speed networks under the Global Interconnection of Broadband Networks programme. NSF also intends shortly to release proposals for an international connections programme, setting out plans for connection points, cooperation and cost-sharing.

This flurry of activity testifies to growing awareness that the linking together of the many networks making up the Internet has been badly neglected. Although many countries have high-performance research networks, links between them are often extremely slow. Once scientists leave the 'motorway' of their own national network, the only routes to other national motorways often resemble dirt tracks.

Scientists are particularly badly affected, because they require the highest Internet performance for advanced collaboration techniques (see *Nature* 380, 377; 1996). "Science is a global activity and US scientists need access to facilities in Europe and Japan. We need to look at [Internet development] as

a global enterprise, and not as a national enterprise," says George Strawn, NSF's division director for networking and communications research and infrastructure.

Many European countries have national research networks with bandwidths of 34 Mbps, and most are in the process of moving to 155 and 622 Mbps. But the speed of links between them has often been as low as 2 Mbps. The first Ten-34 links opened last month between research networks in France, Germany, the United Kingdom and Switzerland, and those among other European countries will follow shortly.

The Ten-34 project costs around ECU40 million (US\$45 million) annually, with ECU12 million coming from the European Union. It is run by the Ten-34 consortium, a group of telecommunications operators and 16 national research networks, led by Dante, a not-for-profit company based in Cambridge, England, which was established in 1993 by national university networks.

Similar problems of interconnectivity occur in the United States. Since the privatization of the NSFnet in 1994, universities can connect with each other only across the crowded public Internet. Although the performance of the public system is improving, it often falls short of that required for demanding research applications.

Intranets or public Internet?

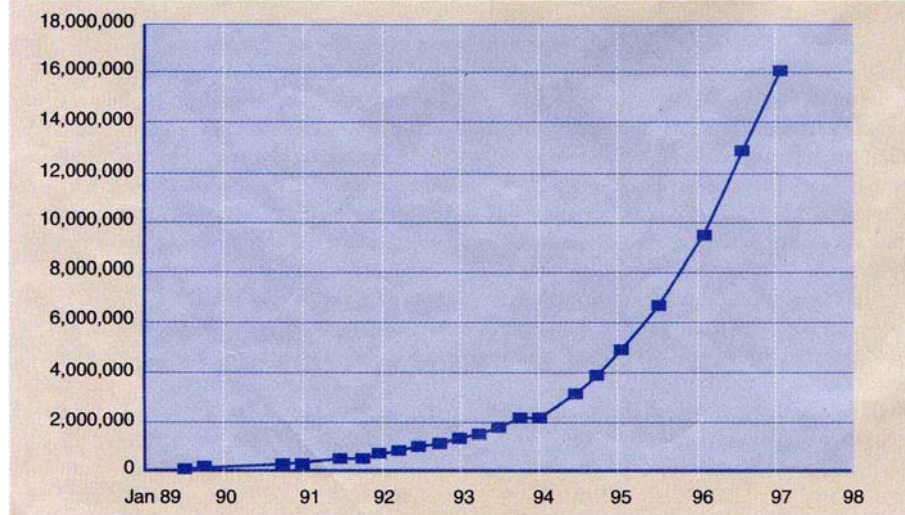
The situation is slightly better for the national laboratories belonging to the federal agencies, such as the Department of Energy, which have powerful 'intranets', cut off from, but connecting to, the public Internet. But connections between the individual agencies are poor, while links between national laboratories and universities have to rely on the public Internet.

"High-performance interconnectivity between individual networks is now a major issue," says Mike Roberts, director of the Internet-2 project. More than 100 universities have now committed \$50 million of their own funds to this project, which also has the backing of major computer companies.

Internet-2's initial goal is to build high-speed links — 100 times faster than today's — among universities and government laboratories. Participants have agreed to build shared regional infrastructure, called gigapops ("gigabit capacity point of presence"), which will interconnect with regional institutions and each other.

Internet-2's applications steering group says it plans to create tools such as multicasting, streaming, application-sharing, and synchronous communications technologies to let users build new distance learning and instructional management systems, and digital library, virtual lab, and 'teleimmersion' applications.

But the ultimate aim of all these initiatives is to act as a testbed for the development



Growth of Internet hosts, 1989–1997. Data available on <http://www.nw.com>.

of high-performance applications, connections and protocols that would transform the public Internet into a high-speed network. This marks a return to a more interventionist approach to Internet development in the United States.

The privatization of NSFnet led to a slip in Internet development, Strawn concedes. There is now a growing realization that the state has a role to play, he says. "People have migrated towards the view that networking advances could be accelerated by reinvigoration of the public-private partnership that brought us the Internet in the first place."

Major benefits for researchers

The United States does not intend to create private research intranets closed off from the public Internet, as still happens in Europe and Japan. Although the vBNS is for the moment a private intranet, it will provide the research community with high levels of performance until these can be delivered by the public Internet, says Strawn. "There is no intention of setting up a separate infrastructure [in the long term]."

The costs of maintaining intranets are exorbitant. Many European research organizations now believe the best performance at lowest cost will eventually come from the public Internet, which will quickly evolve to

high speeds while delivering economies of scale (see *Nature* 378, 380; 1996).

Scott Bradner, a senior technical consultant at Harvard University, agrees, pointing out that researchers have not needed to build an independent telephone network. "There is no long-term viable solution which excludes the public Internet," he says.

Major benefits to scientists will come from a new set of protocols being tested in second-generation Internet projects. These will enable users to select the level of performance they want, with a sliding scale of charges replacing the system where most users pay a flat monthly connection fee.

The idea of prioritizing traffic in this way was until recently opposed by many Internet users attached to the principle of flat-fee access. But it is now gaining widespread acceptance as users grow to realize that popular video applications, like many research applications, will require guaranteed levels of performance if they are to work properly.

The new protocols may also ease the current political problem of sharing the high costs of international links. They would allow countries to be guaranteed a specific share of the bandwidth corresponding to the amount of its subscription, which is impossible at present. Bottlenecks on transcontinental links are a major problem.

Concern about global connectivity recently prompted the International Committee for Future Accelerators — which promotes collaboration in the building and use of accelerators — to set up an ad hoc panel to address the issue. "This is the first time that ICFA have recognized [Internet connectivity] as a serious issue," says David Williams, chair of the panel and former head of the computing and networks division of the European Laboratory for Particle Physics.

Progress in establishing better international links has been hampered in particular by the political difficulties of negotiating them. European research organizations, for example, resent paying 90 per cent of the costs of the shared links to the United States (with NSF paying the rest), whereas each accounts for about the same amount of traffic over the link.

Some of this imbalance is because the cost of bandwidth in European countries where telecommunications companies still have monopolies is an order of magnitude higher than in the United States. Indeed, the high cost of bandwidth remains the major obstacle to general development of the Internet in Europe, although many believe that the scheduled deregulation of telecommunications companies within the union next year will bring about major price reductions.

But Dai Davies, general manager of Dante, attributes much of the imbalance to Europe's relatively weak negotiating position in the past. "Europe does not have a single voice on this, and the Americans have been able to divide and conquer." He says there is also an "asymmetry" in that all European countries have state-supported research intranets whereas Internet services in the United States are mainly supplied by commercial providers which are mainly concerned with local connectivity and "don't care about research links to Europe".

Competition sparks alliances

France, Germany, Italy, the United Kingdom and the European Commission have adopted a common position for this month's meeting on the proposed G7 Global Interconnection of Broadband Networks project. Europe will for the first time have a single voice on the issue, says Davies. He predicts that the Americans "are going to get pinned down" for not contributing sufficiently on global connectivity.

Poor international and regional connections have also resulted from the reluctance of telecommunications operators to link up with each others' networks. Global competition is now prompting a spate of joint ventures — such as Global One, a joint network between France Telecom, Deutsche Telecom and Sprint — which will bring faster and more reliable connections. Mergers and alliances will be extremely important for connectivity", says Strawn. □

'Europe must do more to promote networks'

[PARIS] The European Union (EU) should play a much greater role in promoting both electronic networks for researchers and the public Internet, says a report by one of the European Commission's advisory groups. "This is a plea for recognition [of the importance of the Internet] at the highest levels of the union," says the report. "Europe has to wake up to the Internet at many levels."

The group's members are communications experts from universities, government research organizations and private industry. It was set up last year in response to the creation of several second-generation Internet initiatives in the United States. Its report, *The Future of the Internet – What Role for Europe?*, says there is a need for "urgent and effective action, and appropriate budgetary commitments" to catch up with the United States.

The commission also needs a dedicated body to deal with the Internet, according to the report, which points out that at present "each major Internet development made outside Europe has to be analysed and reacted to by fairly *ad hoc* groups of EU staff and outside experts". The group says it is "urgent" that the commission should appoint experts to track the progress of the US initiatives, which it argues represent a significant development.

The report says that "excessive line costs are still the biggest barrier to the spread of the Internet in Europe". It calls on the EU to act quickly to ensure that the liberalization of European telecommunications companies scheduled for next year results in a fall in bandwidth costs. This will be critical if Europe is to move towards the high-speed networks envisaged in the US second-generation

Internet project, says the report. High speeds are needed by highly demanding applications.

It is "essential" that the EU makes funding available for a follow-up to the Ten-34 project to bring links between national research networks up to 155 and 622 Mbps, the report says. It also recommends that European countries build shared regional infrastructure, along the lines of the gigapops planned in the US Internet-2 initiative (see above).

The wider thrust of the report is that the EU should play a greater role in supporting Internet research and development and infrastructure more generally. Areas where the union could help include supporting the development of Internet protocols for prioritizing traffic and encouraging the deployment of alternative technologies for home access to the Internet, such as cable television and satellite. **D.B.**

Montagnier to set up an AIDS research centre in New York

[PARIS] Luc Montagnier, a member of the group at France's Institut Pasteur that discovered the AIDS virus, is to set up an AIDS and emerging diseases research centre at Queen's College, part of the City University of New York. "It's a very exciting project for me," says Montagnier. Bernard Salick, a doctor and businessman who attended Queen's College, has donated \$4.5 million towards the centre. Montagnier says he hopes to raise \$30 million from other private sources and New York state.

Montagnier expects the centre to be built within two years. It will be associated with the World Foundation for AIDS Research and Prevention, a body created in 1993 by Montagnier and Federico Mayor, director general of the United Nations Educational, Scientific and Cultural Organization. The foundation already has centres in Abidjan in the Ivory Coast and Paris.

The new post of emeritus professor created for Montagnier would require him to spend nine months of the year in the United States. But he says that he will nonetheless try to keep his Paris laboratory until he retires from the Institut Pasteur in three years, "if the institute agrees".

Icelandic biotechnology looks to Vikings

[LONDON] Scientists and doctors in Iceland have raised US\$12 million to set up the country's first biotechnology company. DeCode Genetics will look for genes involved in common multigenic diseases, taking advantage of Iceland's unusually homogeneous population — most of its 300,000 citizens are descended from seventh-century Vikings.

The company is financed by US and European venture capital funds, although Icelanders own most of the shares and make up 52 out of the company's 55 employees. The founders say they are proud of the company's local roots, as Icelanders would be upset if foreigners were found to be exploiting their country's genetic resources.

NASA narrows down mission proposals

[WASHINGTON] Spacecraft missions to study the planet Mercury, the atmosphere of Venus, the moons of Mars, comets and the solar wind were selected for further study last week from 34 proposals submitted to the US National Aeronautics and Space Administration (NASA) last December. Each winner will receive \$350,000 for a four-month feasibility study, after which NASA

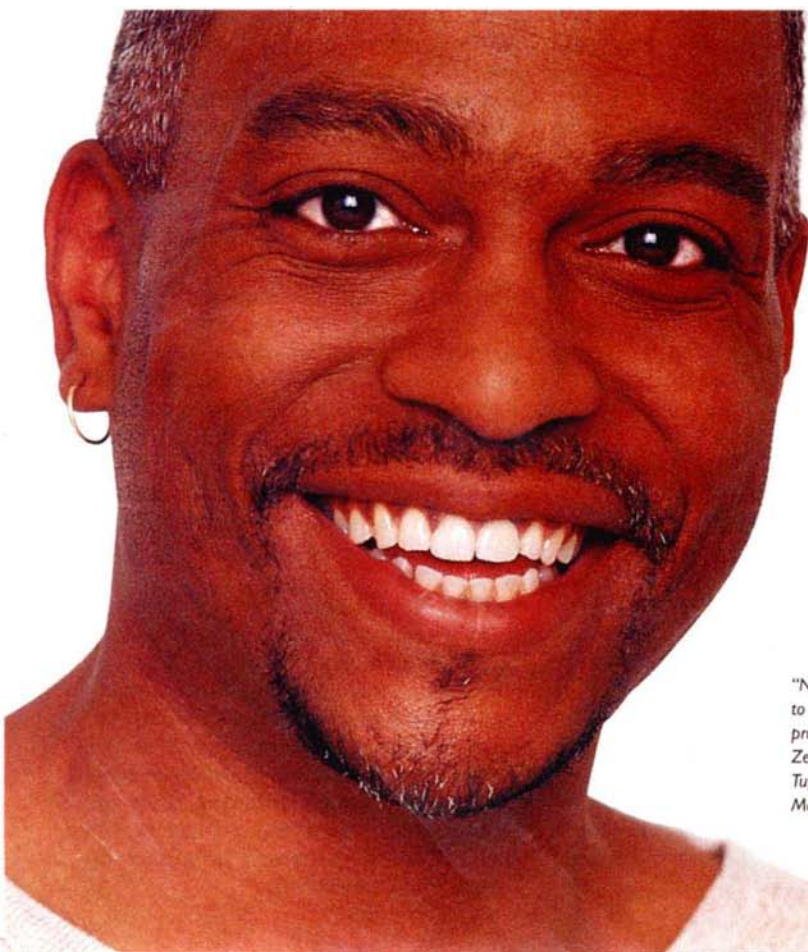
will select one or two for launch as 'Discovery' planetary missions no later than September 2002. The concepts range in cost from \$135 million for the Contour comet nucleus tour, proposed by Joseph Veverka of Cornell University, to \$257 million for a spacecraft proposed by Sean Solomon of the Carnegie Institution to orbit Mercury.

French view of science unmoved by scandals

[PARIS] A French opinion poll has suggested that recent scandals — from the contaminated blood affair to the 'mad cow' crisis — appear to have had little impact on the public's esteem for science.

Exactly half of those questioned think that science does more good than harm, down from 56 per cent in a poll in 1972, but up from 41 per cent in 1989. Some 43 per cent said they had greater confidence in science than a decade ago, compared with just 17 per cent who said they had less. When questioned about responsibility for the recent scandals, just 20 per cent believed that scientists had influenced politicians; 70 per cent thought politicians had used scientists. But three-quarters of those polled thought that the knowledge of scientists could make them dangerous — a figure that is unchanged from 1989.

The poll was commissioned from the



Malcolm is imp
by the new Pre
Protease be
it's a GST
fusion prot

"Not being a protein chemist, I just want to clone the gene, express it, isolate the protein and move on," says Malcolm Zellars, who's working on his post-doc at Tufts University Medical School in Boston, Massachusetts, USA.

pollster Sofres by the magazine *Sciences et Avenir* to mark its fiftieth anniversary, and is based on a sample of 1,000 people.

Joint operation to study Chinese fossil site

[LONDON] Chinese and Western palaeontologists have launched a US\$1-million collaboration to hunt for dinosaur and bird fossils at an extensive, newly found fossil site in northeast China, in a venture sponsored by the Academy of Natural Sciences in Philadelphia. The fossil site, believed to have been found by Chinese farmers, is understood to contain hundreds of specimens of early birds, dinosaurs and insects, all of which are remarkably well preserved. It has already produced two fossils, including *Sinosauropteryx*, a dinosaur that appears to have had a line of feather-like structures running along its spine. Claims that *Sinosauropteryx* was a 'feathered' dinosaur continue to generate sharp divisions among palaeontologists.

Lyme disease expert faces suspension

[WASHINGTON] An expert on Lyme disease at the US National Institute of Allergy and Infectious Diseases (NIAID) may be suspended without pay for two weeks for his

outspoken criticism of an advocacy group. Edward McSweeney, a scientist in the division of microbiology and infectious diseases, recently received a notice of proposed suspension, after NIAID found that his private home page on the World Wide Web contained a link to the 'Whacko Lyme Disease Foundation'.

McSweeney has for years been sharply critical of the foundation, based in Hartford, Connecticut, arguing that it exerts excessive influence on NIAID's research programme. In 1995, he was barred by NIAID from official activity on the disease, but he has continued his crusade privately.

Former *Nature* editor wins cultural award

[MUNICH] Sir John Maddox, the former editor of *Nature*, has been awarded the Eduard Rhein Foundation 1997 Cultural Award "in recognition of his long-standing commitment to furthering and disseminating scientific knowledge as the editor of *Nature*". Worth DM200,000 (US\$130,000), the prize is one of the most valuable cultural awards in Europe. The foundation was set up in 1976 and is named after the German writer and inventor Eduard Rhein, who died in 1993. It uses awards to promote scientific research, the arts and culture in Germany and abroad.

Beekeepers save 'radar bee' tracking project



[LONDON] A £6,000 (US\$3,700) grant from British beekeepers will save a novel research project from closing down and relocating to the United States after it was turned down for funding in the United Kingdom. The research involves tracking the movement of honeybees by gluing tiny radar antennas to their backs (pictured above).

The technique was successfully demonstrated last January (see *Nature* 379, 29; 1996) by scientists at the UK government's Institute of Arable Crops Research in Rothamsted in the south of England. The grant from members of the British Beekeepers Association will support experiments for one year.

ressed
Scission
cause
ein

Are you working with GST fusion proteins? A new one-step solution offers you complete simplicity in purification. Introducing pGEX-6P vectors and a recombinant GST fusion protein called PreScission™ Protease.

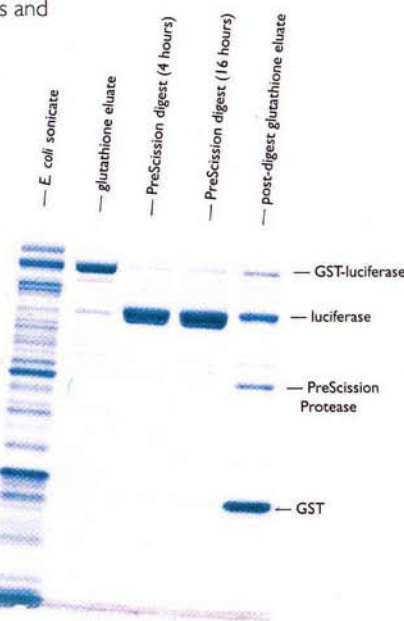
One-step separation of the protease and GST affinity tag from your protein of interest

Following the site-specific cleavage of GST fusion proteins, purification of the protein free of PreScission Protease, and the liberated GST affinity handle, are accomplished in one step while using the same affinity medium, Glutathione Sepharose®.

The revolutionary PreScission Protease is similar to other commercial site-specific proteases in that its recognition sequence is rarely present in most native proteins. What's more, the protease performs at a 5°C temperature optimum to preserve the integrity of your protein. This quick and easy, single-step solution is only available from one company, Pharmacia Biotech.

Are you doing gene expression?

Give us a call: 1 (800) 526-3593 in the USA; 03 3492 6949 in Japan; +46 18 16 50 11 in Europe and the rest of the world. Or visit us on the Internet: <http://www.biotech.pharmacia.se>.



Pharmacia Biotech
Uppsala, Sweden. (And the rest of the world)

Star-gazing funds should come down to Earth

Sir — In January I reported¹ on the results of an analysis I had undertaken on the cost-effectiveness of ground-based optical telescopes in the period 1978–94. I have now extended this analysis to include radiotelescopes and space-based observatories to see if a change is indicated in the balance of future investments in such astronomical observational facilities. As far as I am aware, this is the first such comprehensive cost-benefit analysis undertaken on either side of the Atlantic.

I have assessed the benefit or effectiveness of a particular facility by analysing the 15 per cent most cited papers published in *Astrophysical Journal* and *Monthly Notices of the Royal Astronomical Society* every half-year from 1958 to 1994 inclusive, at four-yearly intervals, and deducing which facilities were used to produce the new observations analysed in those papers. In many cases, more than one telescope or spacecraft were used to produce the data analysed in a given paper and, in that case, a scoring system was used² to take into account these different facilities, the total score for each paper being unity.

I have limited my analysis to observational facilities owned by governments and other institutions in the United States and British Commonwealth but, in the case of spacecraft, I have extended the analysis to include the European Space Agency, as that is where the United Kingdom spends most of its space-related funds.

My analysis of the total costs of these various facilities over the period 1956–92 shows that the balance of funding between capital and annual operational expenditure is very different for ground- and space-based facilities, as shown in Table 1. On this basis, extra capital expenditure on ground-based observatories would be money well spent, if that could significantly reduce the annual operations costs. In the case of

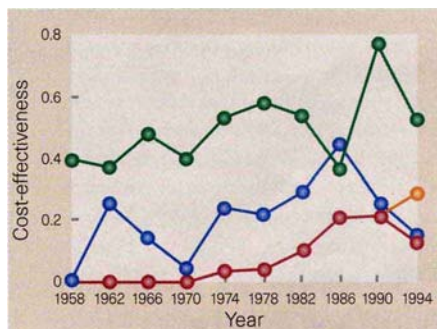


Figure 1 Plots of the cost-effectiveness as a function of time for ground-based optical/infrared and radio observatories and spacecraft, where the cost-effectiveness is defined as the number of highly cited papers per year divided by total annual costs in millions of 1992 dollars and where the total annual costs = annual operations costs + amortized capital costs. Optical is green, radio blue, spacecraft purple. The cost-effectiveness of spacecraft excluding the Hubble Space Telescope is shown in orange.

spacecraft, however, the situation is very different, and a major effort should be made to reduce the capital costs of such facilities. Alternatively, spacecraft should be designed for longer operational lifetimes, if this can be accomplished with only a small increase in capital costs.

Putting my analysis of costs and benefits together produces the cost-benefit results shown in Figure 1. This shows that there is a good case for increasing the investment in ground-based optical/infrared facilities at the expense of space-based astronomical observatories. Although the Hubble Space Telescope has been extremely successful, Figure 1 shows that it has not yet justified its high costs (both capital and operational), even though the relatively long planned lifetime (for spacecraft) of 15 years has reduced the annual write-off of capital costs. The US National Aeronautics and Space Administration's aim of producing a replacement facility in the next decade for about a quarter of the capital cost is the correct approach, provided the annual operations costs can be similarly reduced.

David Leverington

Physics Department,
Open University,
Walton Hall,
Milton Keynes MK7 6AA, UK
e-mail: D.Leverington@open.ac.uk

1. Leverington, D. *Nature* **385**, 196 (1997).
2. Leverington, D. Q. J. R. Astr. Soc. **37**, 643 (1996).

The one true faith

Sir — The Commentary article “Scientists are still keeping the faith” by Edward J. Larson and Larry Witham was superimposed on a large cross occupying more than half the height of the page (*Nature* **386**, 435; 1997). Leuba's survey contained no reference to Christianity and could apply to Judaism or Islam, to name but two other possibilities. Contrary to the belief of many monotheists, there is not a monopoly on monotheism, and I assume that one does not have to sign an undertaking that one does not have a non-Christian background to be included in *American Men and Women of Science*.

Mark R. Baker

Visual Sciences,
Radcliffe Infirmary,
Oxford OX2 6HE, UK
e-mail: mark.baker@green.ox.ac.uk

No colour slides please

Sir — Researchers used to spend days composing drawings to make slides for their dissertations, but most of the problems were overcome when computer programs arrived. Laser printers and plotters made all the drafts before impeccable graphs emerged. The only remaining thing to do was take a photograph. Speakers talked to a wide-awake audience whose members could see each other and take notes.

But now everything has changed. New hardware offers the possibility of making slide pictures directly from the computer and, worse, of applying colour to black-and-white slides.

Presentations now start with a request for the lights to be switched off, and the speaker shows with satisfaction a slide where the conference title has been written in red, on a dark blue background, with a round corner frame in green and with the authors' names in bright yellow. People try to read the text from the abstract book and take notes without knowledge of Braille. Some participants will take advantage of the dark to enjoy a siesta. In a decade in which handicapped people have forced governments to remove physical barriers, colour-blind scientists are losing a battle with their own colleagues. I wonder whether such colourful slides are necessary and what their contribution, if any, is to scientific communication.

Ricardo Borges

Unidad de Farmacología,
Facultad de Medicina,
Universidad de La Laguna,
Tenerife, Spain

Table 1 Costs in 1992 US\$ (million) 1956–92

Observatories	Total annual operations costs	Written-off capital costs
Ground-based optical/IR	3,300 ± 500	210 ± 70
Ground-based radio	2,100 ± 500	580 ± 60
Space-based	1,600 ± 200	4,600 ± 500

The period 1956–72 for the above costs is two years in advance of the period for publication of papers. This is to allow time for the observational data to be analysed, written up and published.

Are patents and research compatible?

At a time when the number of biotechnology-based patent applications is soaring, it is essential that the law allows a significant 'experimental-use' exemption. The United States urgently needs to modify its legislation in this area.

Philippe Ducor

The patent system is generally agreed to promote innovation, by rewarding inventors with a limited exclusivity on their discoveries. But this goal is attained at the expense of subsequent innovators, who must first obtain a licence to use the patented technology before bringing it to the next step.

To avoid a chilling effect on subsequent innovation and to preserve the original purpose of patents, most countries' legal systems allow some degree of 'experimental use' on a patented invention without considering the patent to be infringed. In fields such as biotechnology, where the number of patent applications is soaring but where subsequent research is usually required before society can benefit from specific patented inventions, it is essential to have a significant experimental-use exemption enshrined in law. Although most European countries and Japan have included this principle in their patent statutes, the US Patent Act has not, the system relying on outdated case law. Considering its position as a world leader in biotechnology, the United States should urgently reconsider its policy.

Broadening uses of patents

Until recently, most patented innovations were incorporated in finished goods or other products and processes reasonably close to the market, as only then were they consid-

ered to be worth the cost involved in a patent application. Such inventions typically resulted from corporate scientists or engineers doing product-oriented, 'applied' research, for example the instant camera and the diesel engine. Academics were considered to be doing 'basic' research, motivated by more ethereal goals generally unrelated to financial profit: developing the theory of relativity or discovering the double-helical structure of DNA, to take two famous examples. These worlds remained essentially apart, and patents belonged almost exclusively to the business world.

Patent infringement cases involving experimental use were typically between commercial competitors over marketed or near-market products, and were usually decided against the defendant as the analysis focused primarily on the commercial versus non-commercial intent of the alleged infringer¹. The law, aiming to protect patent owners who would by definition lose no income from a non-commercial venture, seemed adequate at a time when basic research was hardly ever concerned with patents. Accordingly, experimental-use exemption has long been a relatively uncontroversial issue, confined to judicial dicta and isolated cases.

The advent of biotechnology has radically changed the picture, by virtually eliminating the traditional distinction between 'basic' and 'applied' research. Nowadays, fairly basic

scientific information previously considered as strictly academic in nature can have a significant economic value in addition to scientific or academic prestige. As a result, both the biotechnology industry and non-profit research institutions have become interested in the same type of information, and are often equally aggressive in patenting their discoveries. Research collaborations between industry and academic institutions are now widespread, and many biotechnology companies have their origins in university-based research. In this context, the traditional test distinguishing the commercial from the non-commercial purpose of experiments has become essentially inoperative.

On the one hand, the non-profit nature of many university research activities is increasingly questionable. On the other, there is no policy justification to prohibit follow-on research occurring in commercial settings while allowing it in non-profit institutions, as the ultimate benefit to society does not depend on where the research is done.

Patents controlling research

The interest of corporations in basic research and of the academic community in patents has affected not only the number of biotechnology patents issued, but also their nature. A growing number of patents is issued on basic technology, where most of the value lies in the research potential leading to subsequent innovations. As a result, entire areas of research are now controlled by patents.

The Cohen-Boyer patent on basic DNA recombinant technology provides an early example. Thanks to the wise licensing policy of non-exclusivity and modest pricing adopted by the owner (Stanford University), the technology was widely distributed and contributed enormously to the subsequent development of commercial biotechnology. However, the outcome would have been radically different if Stanford had chosen instead a restrictive licensing strategy (exclusive basis, high prices), or granted no licences.

More recently, the Agracetus (now Monsanto) patent covering all genetically altered cotton stirred controversy, as the licence fee for using the technology was fixed at \$1 million. This kind of amount is no obstacle for big companies, but it can force smaller ones out of business. Indeed, Mycogen, a company based in San Diego, temporarily abandoned its cotton research programme for this very reason². Out of concern that the Agracetus patent might seriously hamper cotton research, the US Department of Agriculture and another party petitioned the US

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Patent law must be adapted to new situations to preserve its original goal of promoting innovation.

Patent and Trademark Office for re-examination. The issue is still unresolved, pending appeal.

In the Cohen-Boyer and Agracetus examples, one or a few broad patents control an entire research area. In other instances, many narrower patents have the same effect. To take HIV research, for example, patents have been issued on most aspects of the virus, resulting in a dense network of proprietary rights. Wild-type HIV strains, envelope proteins, regulatory genes, fused proteins and defective viruses have been patented in addition to innumerable other HIV-related inventions. This has prompted the US patent office to create a dedicated database, containing thousands of entries, available on the Internet at <http://www.uspto.gov>. Apart from a genuine 'experimental-use' exemption, scientists intending to use any of the patented technologies for their research have to depend on sensible licensing and enforcement policies from the respective patent holders. This amounts to leaving the implementation of a policy matter — promoting follow-on innovation — to private interests.

This problem could be tackled by questioning the wisdom of granting patents on exceedingly broad or basic technology — the US Department of Agriculture's strategy in the cotton patent case — because this practice potentially dwarfs the innovation-promoting goal of the patent system³. Prohibiting patents on such inventions would by the same token virtually resolve the problem of experimental-use exemption, as broad and basic inventions are the most likely to lead to further experimentation.

Unfortunately, existing patent laws are ill-suited for distinguishing between technologies that are too broad or too basic to justify patent protection, and others. They rely for this purpose on the fuzzy notions of 'utility' (equivalent to 'industrial application' in Europe) and 'enablement', which are often dauntingly difficult to apply in biotechnology. Besides these legal difficulties, there is no policy rationale to reward only inventors whose applications are close to the market, and not the discoverers of more basic innovations which often bring greater overall benefits to humanity. Accordingly, all inventions meeting the patentability requirements of patent law — utility, novelty, non-obviousness and enablement — probably deserve protection, notwithstanding their sometimes very broad or basic nature.

Nuanced solution

Instead of prohibiting certain patents, a true, policy-oriented experimental-use exemption offers a nuanced solution, allowing subsequent research to occur while preserving the interests of initial inventors. Experimental activity intended to bring the technology to the next step is freely allowed, whether in commercial or non-commercial settings.

But as no patents need to be prohibited, initial inventors are allowed to reap part of the profits generated by follow-on research once a product (hence actual income) results.

The legal treatment of the experimental-use exemption currently varies from country to country. In most European countries and Japan, the principle has long been accepted and codified in patent statutes. The idea is to grant immunity for research relating to the patented subject matter, the aim being to bring the technology to the next step. Although the provisions existing in European and Japanese laws need to be constantly reinterpreted as new situations and technologies emerge — not every experimentation can be exempted from infringement — they are firmly grounded in their respective legal cultures and generally provide some room for experimental-use exemption in commercial settings⁴.

Nothing similar exists in US law. Most of the experimental-use exemption doctrine is based on isolated dicta in two 1813 Massachusetts cases, focusing almost exclusively on the commercial versus 'philosophical' intent of the defendant. Accordingly, some level of experimental exemption seems to exist in US law, limited to non-profit research and subject to the uncertainties of uncoded jurisprudence. Amounting to a no-cost compulsory licence, the principle of experimental exemption has never been popular in US legal culture, traditionally averse to all forms of expropriation. Indeed, a codification attempt by Congress, the Patent Competitiveness and Technological Innovation Act of 1990, failed miserably.

In its 1984 decision *Roche v. Bolar*⁵, the Court of Appeals for the Federal Circuit cast doubts on the existence of any experimental exemption in commercial settings. The defendant had conducted regulatory testing with the plaintiff's patented drug Dalmane before the expiration of the patent, to obtain approval for the generic version of the drug as soon as possible after expiry. The court, while correctly deciding that such business testing was not exempted from infringement, insisted in an unnecessary dictum that the experimental-use exemption was "truly narrow", and that "Bolar's intended 'experimental' use is solely for business reasons and not for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry".

Although the main holding in *Roche v. Bolar* was legislatively overruled by Congress shortly thereafter (in the Drug Price Competition and Patent Term Restoration Act), the dictum leaves virtually no room for a commercially significant experimental exemption. This was tacitly confirmed in *Scripps v. Genentech*⁶, decided three years later. In this case, Scripps was the owner of a patent covering the human blood-clotting agent factor VIII:C, as well as a purification method to obtain it. Genentech had used the purified

factor to retrieve the human gene encoding it, which at the time was certainly not a routine procedure. Genentech subsequently succeeded in producing recombinant factor VIII:C by cloning the gene, and was about to commercialize it when Scripps sued for patent infringement.

Understandably, most of the attention in the case was drawn to the alleged infringement by Genentech in making recombinant factor VIII:C. However, the issue of experimental use was also very relevant. The district court judge decided without much explanation that the use by Genentech of the patented purified factor VIII:C as a starting point towards the corresponding gene infringed Scripps' patent. Genentech itself did not invoke the experimental-use exemption, and the issue was not raised by either of the two courts that subsequently reviewed the case.

However, what Genentech did at the time with purified factor VIII:C was precisely the type of activity that should fall under an experimental exemption: using existing technology to bring it to the next step. Purified factor VIII:C was not only an end-product with a significant market; it was also a research tool indispensable to the next technological step. Although Genentech's production of recombinant factor VIII:C was arguably infringing Scripps' patent — commercialization of research results must be distinguished from research itself — the initial experimental use of the purified factor VIII:C should have been exempted from infringement. It is revealing that neither the successive courts in charge of the case nor the defendant itself made the distinction.

Many patents in biotechnology today are research tools (invitations to take the next technological step) rather than traditional end-products. In this context a true, policy-oriented experimental-use exemption is essential if the ultimate goal of the patent system is to be preserved. Failing to adapt to new research patterns, current US patent law threatens this goal by allowing research pertaining to specific fields to be confined to the relevant patent owners, and so solutions beyond the owners' competence or imagination can be overlooked forever. That was not the intent of those who drafted the Patent Act. □

Philippe Ducor is at Stanford Law School, Stanford, California 94305-8610, USA. Correspondence address: 18 chemin Pré Cartelier, 1202 Geneva, Switzerland.
e-mail: philduke@leland.stanford.edu

1. Bruzzone, L. *Amer. Intell. Property Law Assoc. Q. J.* 21, 52-69 (1993).
2. Stone, R. *Science* 268, 656-658 (1995).
3. Eisenberg, R. *Risk: Health, Safety and Environment* 5, 163-175 (1994).
4. Parker, D. *Houston J. Int. Law* 16, 615-663 (1994).
5. *Roche Products v. Bolar Pharmaceutical Co.* 733 F.2d 858 (Federal Circuit, 1984).
6. *Scripps Clinic and Research Foundation v. Genentech Inc.* 666 F. suppl. 1379 (Northern District of California, 1987).

Natural selection out on a limb

Ted J. Case

Patience is a virtue for the evolutionary biologist. Long-term experiments on populations of lizards introduced onto islands in the Bahamas show the forces of adaptation to environment at work.

One criticism sometimes levelled at evolutionary biology is that there are few well-documented cases in which evolution by natural selection has actually been witnessed outside the laboratory. This is particularly true for vertebrate species: the burden of proof is heavy when the phenomenon concerned occurs on a timescale typically longer than a human lifetime. Yet, if natural selection is strong enough, then its adaptive results might be manifested in relatively short order.

Recently colonized islands provide the necessary ingredients for rapid evolutionary change — a small initial population size, a water barrier that diminishes gene exchange with other areas, and different selective regimes. Last year, for example, Cody and Overton¹ reported the rapid loss of dispersal ability in wind-dispersed weedy plants in the daisy family (Asteraceae) on small islands off British Columbia. On newly colonized

islands, they witnessed an evolutionary enlargement in the embryonic portion of the seeds and a reduction in the size of the parachute-like pappus that keeps the seeds aloft. These changes are adaptive because they reduce dispersal, which on such tiny islands often results in seeds being lost in the ocean.

On page 70 of this issue² Losos, Warheit and Schoener now describe the remarkable results of a 10–14-year experiment with *Anolis* lizards experimentally introduced onto 14 very small Bahamian islands. All came from a nearby source population on the small island of Staniel Cay, Exumas, Bahamas. The experimental islands lacked lizards of any kind; the species introduced was *Anolis sagrei* (Fig. 1), a relatively good colonist as anoles go. It is thought to have evolved in Cuba, and subsequently dispersed to Jamaica and the Bahamas, and more recently to southern Florida^{3,4}.

The original experimental introductions began in 1977, the object being to determine the effect of island size and founder numbers on colonization success⁵. Some of the original islands carried less than 1 m² of vegetated area whereas others had as much as 6,000 m² or so. The islands were stocked with either five or ten individuals and the fate of the lizards was followed over the next five years.

Populations seeded on the smallest islands were immediate failures; the lizards disappeared within weeks. Populations on most larger islands were successes, and thrived regardless of the initial number of individuals. Losos *et al.*² now report on the longer-term morphological changes in lizards on 11 of these islands plus three new islands experimentally seeded with *Anolis sagrei* in 1981.

What, then, happens when you take anoles, place them on several islands without large trees, wait 10–14 years, and measure the morphological differences between them? The island populations have differentiated, particularly in relative hindlimb length, to become closer to the 'optimal' phenotype to be expected given the shorter and thinner vegetation on their new homes (Fig. 2). Moreover, the degree of reduction in relative hindlimb length across the 14 experimental islands parallels the extent to which the local vegetation departs from that in their common source population on Staniel Cay.

In previous studies, Losos and others documented the striking associations between anole habits and morphologies, and the adaptive nature of these different morphologies in terms of performance measures such as running ability on different substrates^{6–9}. Anoles that perch on narrow branches and twigs are smaller and have shorter hindlimbs than those that perch on wider branches or tree trunks. With wide-diameter perches, speed is in more demand, and longer limbs provide a speed boost at the expense of reduced agility.

On the large islands of the Greater Antilles, different anole species are distributed according to perch type, body size and typical climatic conditions. Out of the essentially infinite number of ways of partitioning these three habitat dimensions and combining them with different morphological



Figure 1 Subject to experiment — the lizard *Anolis sagrei*.



Figure 2 An example of the vegetation on the islands onto which *A. sagrei* was introduced. Notably, large trees are absent.

J. B. LOSOS

T. W. SCHOENER

and colour schemes, anoles of the Greater Antilles seem to have settled on a handful of 'ecomorphs', which have parallels on the different islands. This is remarkable because it is believed that the ecomorphs are convergent between islands⁶. Thus, for example, the 'tree trunk-and-crown' ecomorph on Puerto Rico is more closely related to other ecologically dissimilar ecomorphs on Puerto Rico, than it is to the trunk-and-crown anoles on Jamaica or Hispaniola. In this way there is a crude 'periodic table' of types of anole in the Greater Antilles. To be sure, some odd-ball, 'rare-element' species exist, particularly on the large islands of Cuba and Hispaniola, but students of this group have been struck more by the one-by-one species convergences than the occasional departures from the normal ecomorph series.

This story took an exciting turn in the past few years as more detailed phylogenetic analyses, stemming in part from studies on different enzyme forms and on chromosomes, allowed inferences to be made as to the historical build-up of the ecomorph series on Jamaica and Puerto Rico. Not only do these islands share a similar suite of ecomorphs, but the inferred evolutionary build-up of ecomorphs has apparently followed similar trajectories^{8,10}. Clearly, it would be nice to verify these events with a phylogenetic reconstruction based on DNA sequence data.

One outstanding question about the rapid morphological changes in the experimental *A. sagrei* populations is whether those changes are genetic in origin. Anoles are not ideal laboratory animals for genetic studies, and there is a singular lack of reports of the heritabilities of the relevant morphological characters (although there have been studies of a few other lizards and several of snakes). Common experience teaches us that exercise alone can influence bone and muscle growth (witness the couch potato). Yet, even assuming that differentiation might initially represent a purely phenotypic response to altered 'exercise' conditions, it would be surprising indeed if natural selection did not later reinforce these tendencies — phenotypic plasticity typically has a genetic basis^{11,12} and the changes are clearly adaptive.

The morphological rate of change for lizards on these small islands ranged from 0 to 2,117 darwins (a darwin is the change in the mean of the natural log of the characteristic concerned divided by elapsed time in millions of years; the log transformation scales the phenotypic change to remove a potential bias that favours larger absolute changes for larger phenotypic characters). A rate of 2,000 darwins is several orders of magnitude higher than that seen in the fossil record¹³. Reznick *et al.*¹⁴ have, however, reported even higher rates (3,700 to 45,000 darwins) for some life-history characters of guppies. In these experiments, guppies

from high-predation parts of streams in Trinidad were introduced into upstream areas which larger fish predators could not reach. Over 11 years these guppies evolved a later reproductive maturity which occurred at larger body sizes. Given a trade-off between investing resources into adult survival or reproducing, in the former circumstances the selective scales are tipped towards earlier reproduction, because an adult's chance of surviving is severely limited by predators.

The rapid and clear-cut within-species adaptive differentiation of anole populations on tiny islands mirrors in miniature some elements of the larger-scale, across-species adaptive radiations seen in *Anolis* in the Greater Antilles. Yet if significant changes in morphology emerge after only a decade, why do we not see *A. sagrei* 'radiating' even more across the hundreds of islands in the Bahamas where it has been present for probably thousands of years? Populations of *A. sagrei* on different islands have differentiated adaptively: those using broader perches have longer hindlimbs⁷. Still, an *A. sagrei* seems to be an *A. sagrei* throughout the islands, though it may differ a bit in hindlimb length. Perhaps the next major leap in adaptive radiations requires the presence of related species to help force divergence. The

Bahamas has several anoles, although none is endemic. So this ingredient seems to exist, but perhaps these other species are already too different from *A. sagrei* to provide a competitive catalyst for further change. Given the rapid and convincing response reported by Losos *et al.*, it may well be feasible to devise experiments to test interspecific competition as an additional selective force. But the virtue of patience will again be required. □

Ted J. Case is in the Department of Biology, 0116, University of California at San Diego, 9500 Gilman Drive, La Jolla, California 92093-0116, USA.

1. Cody, M. L. & Overton, J. *J. Ecol.* **84**, 53–62 (1996).
2. Losos, J. B., Warheit, K. I. & Schoener, T. W. *Nature* **387**, 70–73 (1997).
3. Wilson, L. D. & Porras, L. *Univ. Kansas Mus. Nat. Hist. Spec. Publ. No. 9* (1983).
4. Salzberg, M. A. *Ecology* **65**, 14–19 (1984).
5. Schoener, T. W. & Schoener, A. *Nature* **302**, 332–334 (1983).
6. Williams, E. E. in *Lizard Ecology: Studies of a Model Organism* (eds Huey, R. B., Pianka, E. R. & Schoener, T. W.) 326–370 (Harvard Univ. Press, Cambridge, MA, 1983).
7. Losos, J. B. & Sinervo, B. *J. Exp. Biol.* **145**, 23–30 (1989).
8. Losos, J. B. *Syst. Biol.* **41**, 403–420 (1992).
9. Losos, J. B., Irschick, D. J. & Schoener, T. W. *Evolution* **48**, 1786–1798 (1994).
10. Losos, J. B. *Phil. Trans. R. Soc. Lond.* **349**, 69–73 (1995).
11. Waddington, C. H. *The Strategy of the Genes* (Allen & Unwin, London, 1957).
12. West-Eberhard, M. J. *Annu. Rev. Ecol. Syst.* **20**, 249–278 (1989).
13. Gingerich, P. D. *Science* **222**, 159–161 (1983).
14. Reznick, D. N., Shaw, F. H., Rodd, F. H. & Shaw, R. G. *Science* **275**, 1934–1937 (1997).

Transcriptional control

Sinful repression

Alan P. Wolffe

The structure of chromatin — the material of which eukaryotic chromosomes are composed — presents both problems and opportunities to the transcriptional machinery. Genes are transcribed in the eukaryotic nucleus without apparent impediment, even though the DNA template is wrapped around histone proteins to form nucleosomes and the chromatin fibre. Evolution has been remarkably successful in shaping chromatin such that it can become alternatively transparent or opaque, to facilitate or restrict the access of transcription factors and RNA polymerase to DNA. The molecular mechanisms that control this access are central to gene regulation, and two reports, by Heinzel *et al.*¹ and Alland *et al.*² on pages 43 and 49 of this issue respectively, along with five related papers in *Cell*, complete one picture of how gene regulation is controlled.

Nucleosomes are not static entities — they adopt a variety of stability states that are dependent on post-translational modification of the core histones. In particular, acetylation of the histones can destabilize nucleosomes and relieve transcriptional repression by allowing transcription factors access to recognition elements. Conversely, deacety-

lation of the histones stabilizes the repressed state³. Last year, certain transcriptional co-activators were found to have histone acetyltransferase activity. So specific targeting of these regulatory complexes could allow nucleosome modification to be directed to particular promoters^{4,5}.

The CREB-binding protein (p300/CBP) is one such co-activator. It has histone acetyltransferase activity, integrating transcriptional-activation signals from diverse regulatory proteins, including steroid and nuclear-hormone receptors⁶. Histone acetylation is targeted by these receptors in response to the addition of ligand (for example, oestrogen and thyroid hormone), rendering the chromatin much more transparent to the transcriptional machinery. Heinzel *et al.*¹ and Alland *et al.*² now show that histone deacetylation might also be targeted by nuclear-hormone receptors and other transcriptional repressors. The five papers in *Cell* substantiate these observations, and extend the repertoire of proven and potential regulatory proteins that use similar repression mechanisms^{7–11}. The emerging picture is one of gene regulation in which reversible alterations in histone chemistry have an important controlling function.

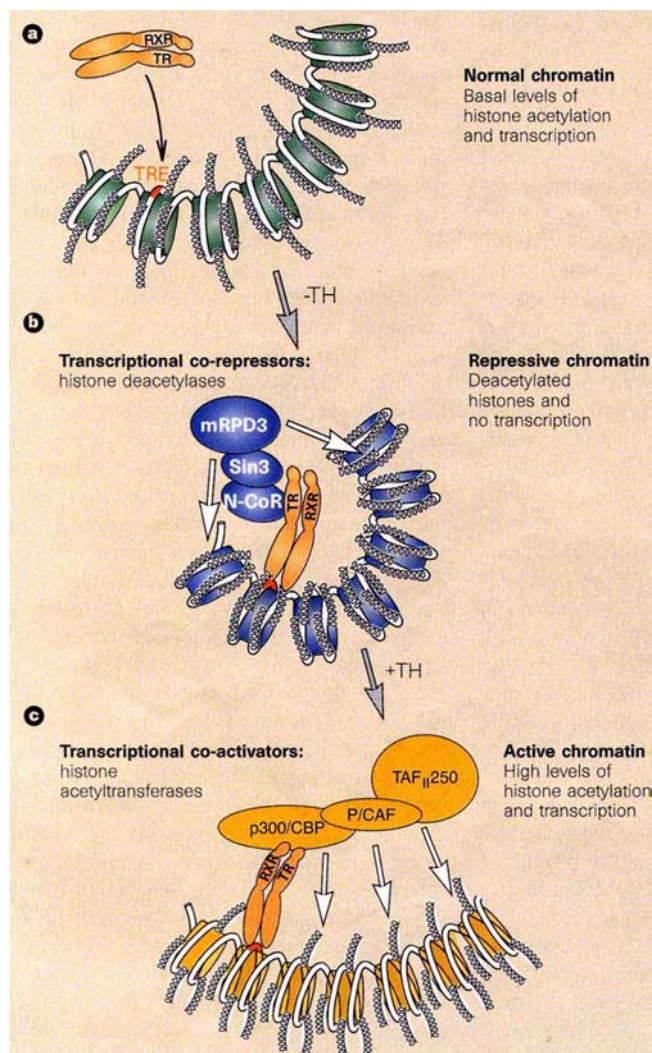


Figure 1 The findings by Heinzel *et al.*¹ and Alland *et al.*² contribute to a three-step model for transcriptional regulation by the thyroid-hormone receptor. **a**, Precise rotational positioning of the DNA sequence containing the thyroid-hormone response element (TRE) on the surface of the histone octamer, allows the thyroid-hormone receptor (TR/RXR) access to chromatin¹⁴. **b**, Once in place, the unliganded receptor recruits a deacetylase complex (mRPD3/Sin3/N-CoR) to augment repression^{1,2}. **c**, Subsequent addition of thyroid hormone (TH) leads to recruitment of acetyltransferases (p300/CBP, P/CAF, TAF_{II}250) that can disrupt these repressive histone-DNA interactions^{5,6,17}.

Heinzel *et al.*¹ report that N-CoR, a known co-repressor for the thyroid-hormone receptor, exists in a complex with the mammalian homologues of two yeast co-repressors, Sin3 and RPD3. Sin3 and RPD3 both act in a pathway that accentuates the transcriptional silencing of several yeast genes¹², and RPD3 is a histone deacetylase¹³. The authors have done antibody-blocking experiments in microinjected mammalian cells to show that each component of the N-CoR/mSin3/mRPD3 complex is essential for the transcriptional repression that is directed by the unliganded thyroid-hormone receptor. Moreover, this transcriptional repression is greatly potentiated by nucleosome assembly¹⁴, in agreement with the prediction that the targeted co-repressor will need a nucleosomal template with which to function most effectively through the action of histone deacetylase.

Mammalian Sin3 is also a co-repressor for the Mad/Max complex¹⁵. Mad is a sequence-specific DNA-binding helix-loop-helix protein that antagonizes the transcriptional activation and transformation functions of the oncoprotein Myc. Mad acts, in part, by competing with Myc for a common heterodimeric partner, Max.

Alland *et al.*² show that mSin3 uses distinct domains to interact with Mad, N-CoR and the mammalian histone deacetylase (mRPD3). This modular series of interactions allows discrimination between the diverse functions of mSin3. Mammalian Sin3 contains four paired, amphipathic helices (PAH1–4). N-CoR associates with PAH1, the Mad-mSin3 association requires PAH2, and mRPD3 associates with the PAH3 and PAH4 domains^{2,15}. Transcriptional repression by Mad requires only PAH1 and PAH2, whereas suppression of transformation requires PAH3 and PAH4 (ref. 2). Histone deacetylase is only recovered in the co-repressor complex that contains all four helices, suggesting that N-CoR bound to PAH1–2 might use other mechanisms — in addition to histone deacetylation — to repress transcription. In contrast, the tumour-suppression properties of mSin3 require the full-length protein and associated histone deacetylase. This is particularly interesting because acute myelocytic leukaemia is associated with a chromosomal translocation that fuses the zinc-finger DNA-binding domain of the MOZ (monocytic-leukaemia zinc-finger) gene to the histone acetyltransferase

CBP (ref. 16). Deregulated histone acetylation might promote transformation, and the regulated activity of histone deacetylase might contribute to growth control.

By defining transcriptional repressors as histone deacetylases and transcriptional activators as histone acetyltransferases, the histones and chromatin structure are fully integrated as necessary components of the regulatory machinery. Although the exact biochemical mechanisms are not yet known, a three-step mechanism can now be proposed for transcriptional regulation by the thyroid-hormone receptor (Fig. 1). Normal chromatin has a constitutive level of histone acetylation that allows basal transcription; but, in the absence of thyroid hormone (TH), the thyroid-hormone receptor (TR/RXR) binds to nucleosomal DNA and targets histone deacetylation, thereby repressing transcription. In the presence of thyroid hormone, histone acetyltransferases are recruited to relieve the transcriptional repression.

Targeted acetylation allows the basal machinery to displace nucleosomes, assemble a functional transcription complex and never have to deal with chromatin again. Alternatively, the regulated association and activity of histone acetyltransferases and deacetylases might occur within a common complex. Transcriptional activity could then be continually modulated through variation in the conformation of the chromatin. Histones would remain present throughout the transcription process. Recapitulating these events means that we will have to understand the regulation of an increasingly sophisticated machinery, and an enzymology that is dedicated to communication between the proteins that package DNA and those that use it as a template. □

Alan P. Wolffe is in the Laboratory of Molecular Embryology, National Institute of Child Health and Human Development, NIH Building 18T, Bethesda, Maryland 20892-5431, USA.

- Heinzel, T. *et al.* *Nature* **387**, 43–48 (1997).
- Alland, L. *et al.* *Nature* **387**, 49–55 (1997).
- Ura, K., Kurumizaka, H., Dimitrov, S., Almouzni, G. & Wolffe, A. P. *EMBO J.* (in the press).
- Brownell, J. E. *et al.* *Cell* **84**, 843–851 (1996).
- Yang, X. J., Ogryzko, V. V., Nishikawa, J. I., Howard, B. & Nakatani, Y. *Nature* **382**, 319–324 (1996).
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. & Nakatani, Y. *Cell* **87**, 953–959 (1996).
- Hassig, C. A., Fleischer, T. C., Billin, A. N., Schreiber, S. L. & Ayer, D. E. *Cell* **89**, 341–347 (1997).
- Laberty, C. D. *et al.* *Cell* **89**, 349–356 (1997).
- Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. *Cell* **89**, 357–364 (1997).
- Kadosh, D. & Struhl, K. *Cell* **89**, 365–371 (1997).
- Nagy, L. *et al.* *Cell* **89**, 373–380 (1997).
- Vidal, M. & Gaber, R. F. *Mol. Cell Biol.* **11**, 6317–6327 (1991).
- Taunton, J., Hassig, C. A. & Schreiber, S. L. *Science* **272**, 408–411 (1996).
- Wong, J., Shi, Y.-B. & Wolffe, A. P. *Genes Dev.* **9**, 2691–2711 (1995).
- Ayer, D. E., Lawrence, Q. A. & Eisenman, R. N. *Cell* **80**, 767–776 (1995).
- Borrow, J. *et al.* *Nature Genet.* **14**, 33–41 (1996).
- Wong, J., Shi, Y.-B. & Wolffe, A. P. *EMBO J.* (in the press).

Superconductivity's cousin

Eduardo Fradkin

The theoretical description of high-temperature superconductors has always assumed a connection between antiferromagnetism and superconductivity. However, the relation of antiferromagnet to superconductor was very similar to the need for scaffolding to construct a building: necessary, but irrelevant to the final result. But according to S. C. Zhang¹, the two phenomena are aspects of the same thing, and can be understood in terms of a new unifying principle based on SO(5) symmetry.

Superconductors lose all electrical resistivity below a critical temperature. In 1957, Bardeen, Cooper and Schrieffer (BCS) developed a theory of superconductivity based on Bose–Einstein condensation of Cooper pairs, bound states of two electrons. This charge condensate is characterized by an order parameter, which determines the local probability of pair creation, and thus has an amplitude and a phase. The phase of quantum-mechanical systems with a fixed number of particles is arbitrary—the number of particles and the phase of the wavefunction are ‘complementary quantities’. Conversely, large systems in a superconducting state have a well-defined phase, and particles can fluctuate in and out of the condensate.

The invariance of the quantum-mechanical hamiltonian under redefinitions of the phase is usually called gauge invariance or U(1) charge symmetry. This symmetry is spontaneously and dynamically broken in a superconducting state. A consequence of the spontaneous symmetry breaking is that phase modulations of the order parameter have an energy, which becomes very small as their wavelength becomes very large. These are called the Goldstone bosons of the superconductor.

In the ground state of an antiferromagnet, the spins of neighbouring atoms are antiparallel, alternating in orientation relative to a fixed but arbitrary vector, the order parameter. The existence of a non-zero order-parameter vector breaks the 3D-rotational, or SO(3), symmetry of the spin system. Here, too, owing to the spontaneous breaking of the spin rotation symmetry, the lowest-energy excitations are Goldstone bosons, known as antiferromagnetic spin waves.

How can such seemingly different states be related? The dynamics of all systems of interacting electrons are compatible with both the U(1) charge symmetry and the SO(3) symmetry of spin rotations. Zhang argues that if the space of allowed symmetry operations on the states of the system is enlarged to SO(5), which includes both U(1) and SO(3), it becomes possible to ‘rotate’ a

superconducting state into an antiferromagnetic state and vice versa. Most systems of interacting electrons do not have the full SO(5) invariance, however, because it is not respected by the dynamics of the electrons and it is violated by explicit terms in the quantum-mechanical hamiltonian, the operator that governs the evolution of states.

The most important symmetry-breaking term is proportional to the chemical potential, and controls the changes in the total charge. Zhang shows that the effect of this term is analogous to the problem of a spinning top in a uniform gravitational field, or a magnetic moment in a uniform magnetic field, as in nuclear magnetic resonance. In the presence of the symmetry-breaking field, the order parameter is forced to precess, and this explicit breaking of the SO(5) invariance induces rotations between superconducting and antiferromagnetic states, and governs the competition between these two phenomena. This precession may already have been detected in neutron scattering experiments.

How is all this related to the physics of high-temperature (high- T_c) superconductors? These materials are usually copper oxides, with other atoms such as lanthanum or yttrium. Their charge carriers are confined to move on two-dimensional planes formed by copper and oxygen atoms. There is an effective magnetic moment with a spin of 1/2 in every unit cell, usually associated with copper. In this picture, oxygen, particularly outside the planes, acts as a charge reservoir, and changes in the oxygen content or ‘doping’ change the total charge of the planes.

Since the discovery of high-temperature superconductors, it has been known that they are poor conductors in their normal state and that minor changes in their structure drive them into insulating, antiferromagnetic states. Short-range antiferromagnetic quantum fluctuations provide the necessary glue to bind the charge carriers into Cooper pairs. But before superconductivity sets in, the antiferromagnetic order must disappear almost completely. In a way, this is like the scaffolding used to build a house: necessary for initial stability but eventually redundant.

Unlike conventional superconductors, the order parameter of high- T_c superconductors is not invariant under spatial rotations. Instead it has d-wave symmetry: it changes sign for every 90-degree rotation in the copper-oxide plane. There is also strong evidence of an energy gap between different spin-carrying excitations, a ‘spin gap’, at temperatures above the onset of superconductivity but below a higher energy scale.

The large size of the spin gap suggests that

it is related to a tendency of the charge carriers to bind into spin singlet states. Interestingly, both superconductivity and antiferromagnetism also result precisely from this tendency to form singlets. At such energy scales it is not possible to determine whether in the low-energy regime the system should be a superconductor or an antiferromagnet. This observation that at some high energy (or temperature) scale it is not possible to differentiate between a superconductor and an antiferromagnet suggests the need for a unifying principle.

Rather than constructing a microscopic theory to explain both superconductivity and antiferromagnetism from a common dynamical principle, Zhang uses a symmetry principle to explain the phases of these systems. By combining the three components of the antiferromagnetic order parameter and the two real components of the d-wave superconducting order parameter into a single five-component vector, he constructs the transformations of the full SO(5) symmetry. He argues that in the microscopic models the full SO(5) symmetry is broken explicitly by the chemical potential while the SO(3) and U(1) symmetries of spin and the U(1) charge remain exact. He shows that, approximately, the microscopic models are consistent with an enlarged symmetry only if the order parameter of the superconductor has d-wave symmetry. Once this crucial assumption is made, a number of known facts now become related to one another in interesting and unsuspected ways.

The requirement of SO(5) symmetry determines the form of theory for the low-energy excitations, and so accounts for the known phase diagram of high- T_c superconductors. Depending on the strength of quantum fluctuations, the ground state can have a direct first-order phase transition, as a function of chemical potential, from an antiferromagnet to a d-wave superconductor; or a quantum disordered spin-gap phase can sneak in between; or there can be a fourth mixed ‘spin bag’ phase.

If instead of the chemical potential, the total charge is fixed, the first-order transition becomes a regime of phase separation. The antiferromagnet and the superconducting phases have the usual Goldstone bosons. However, the SO(5) symmetry reveals itself through the existence of additional low-energy massive modes which would be Goldstone bosons if SO(5) were exact. In the superconducting phase, Zhang identifies these additional modes with the spin-triplet resonance at 41 meV in neutron scattering^{2,3} observed in YBaCuO. And in the antiferromagnetic state, his theory predicts the existence of two massive modes just below the charge energy gap. These modes may already have been detected by Raman scattering.

Another important prediction of this theory is that the core of magnetic vortices

penetrating the superconductor should be antiferromagnetic. This spatial variation should break time-reversal invariance, and so be observable in tunnelling experiments.

Zhang's theory offers a new outlook on high- T_c superconductors. It describes the physics at energies up to the spin gap scale. It does not single out a special dynamical mechanism, but demands that the mechanism must yield a d-wave charge condensate and explain the interplay between antiferromagnetism and superconductivity. There

are several theories that address one or other aspect of this problem, but none that accounts for their complementarity. It is tempting to hope that at low energies these theories will reduce to Zhang's SO(5) theory, thus proving their physical equivalence. □

Eduardo Fradkin is in the Department of Physics, University of Illinois at Urbana-Champaign, 1110 W. Green Street, Urbana, Illinois 61801, USA.

1. Zhang, S. C. *Science* 275, 1089–1096 (1997).
2. Rossat-Mignod, J. et al. *Physica C* 86, 185–189 (1991).
3. Dai, P. et al. *Phys. Rev. Lett.* 77, 3490 (1993).

Field theory

Physicists learn to tie a knot

Warren Perkins

Much of physics is about point-like things — atoms and particles. But extended objects are important in many fields, from solitary waves in hydrodynamics to a whole menagerie of structures in cosmology and condensed matter — sheet-like structures, strings and three-dimensional lumps. Their dimension is important: in cosmology, lumps would cause the Universe to collapse rapidly, whereas strings are a sensible mechanism for seeding large-scale structure formation. On page 58 of this issue¹, Faddeev and Niemi present the first theoretical method of producing lump solitons, by tying string into knots.

The term soliton is applied to any finite-energy, stable, extended object, from the first soliton to be recognized — a non-dispersive 'solitary wave' in a canal — to much more robust objects that are stabilized by their topology. The simplest topological solitons are the kinks in the coiled wire linking a telephone to its handset. These kinks separate regions of left-handed and right-handed coil, and they can only be removed by moving them off the end of the coil or by bringing a kink and an antikink together to annihilate.

Many more complicated defects appear

in condensed-matter systems. In the so-called nematic phase of liquid crystals, for example, the rod-like molecules attempt to align, but this may be prevented if there is a twist in the pattern (Fig. 1). Although the rods are nearly parallel at the edges of the region, a winding of 180 degrees has been introduced which ensures that there is a core region where the orientation of the rods changes rapidly. Extending the pattern to three dimensions produces a linear core region in which the rods are misaligned. This string can move through the material by locally twisting the rods, but it cannot be broken or removed unless it encounters another string and annihilates.

Abstracting these ideas to field theory, the physical molecular rod of the liquid crystal is replaced by a field, and the orientation of the rod is replaced by the orientation of the field. The behaviour of fields is determined by equations of motion (Maxwell's equations, in the case of electromagnetism) which can be obtained from a lagrangian — a function that represents the total energy of the system.

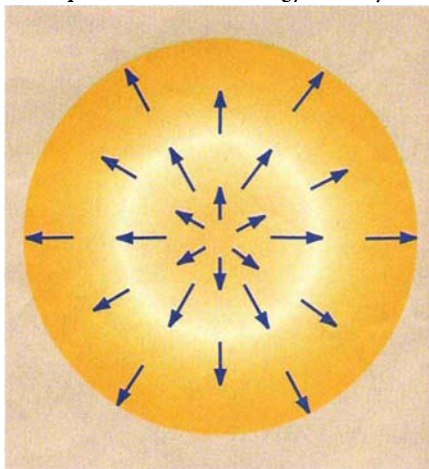


Figure 2 String cross-section. At the central point, the configuration forces this field to go to zero, and change its orientation rapidly. That deviation from the zero-energy configuration (a parallel field of a particular strength) gives the string a mass.

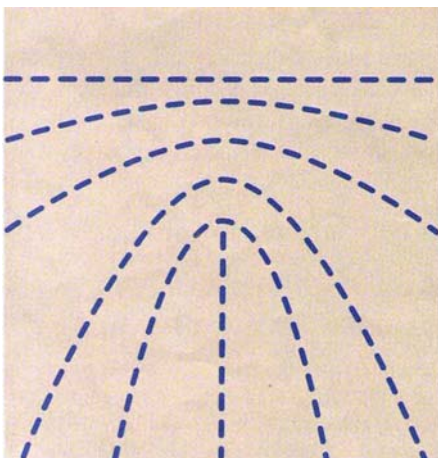


Figure 1 Twisted crystal. The point at which these molecular rods change their orientation rapidly is a stable topological defect.

This determines how the fields evolve over time, or for static field configurations prescribes the energy cost associated with spatial changes in the field, and deviations from any preferred, vacuum, values.

The simplest field-theory string is constructed from a field with two components, which can be represented at any point by an arrow in the x - y plane, whose length is the field's strength. In general, the lowest energy is obtained when all of the arrows are of a specific length and pointing in the same direction. If a winding is introduced (Fig. 2), the arrows remain nearly parallel away from the centre, but there is a core within which the angles change rapidly and the field is forced away from its vacuum value. That, and the spatial gradients in the core, both incur an energy cost; in other words they give the string a mass. The energy cost of moving the field away from its vacuum value ensures that the string cannot be broken, but rather persists unless it encounters another of opposite winding. The mass of the string also results in a tension, which usually causes loops of string to contract, and eventually to disappear.

Massive strings might produce nucleation regions for large-scale structure formation in the Universe, whereas lump solitons, or monopoles, are a cosmological disaster, producing an energy density typically 15 orders of magnitude too high for our Universe still to exist. This problem was the original motivation for the inflationary Universe model, which dilutes the monopoles by introducing a period of accelerated expansion. Similar problems arise if massive string loops can't decay.

Until now no one has been able to construct a knot in field theory, which means that testing theories for knot production has been impossible. Faddeev and Niemi have finally succeeded¹, using a three-component field whose magnitude is fixed, and a lagrangian with two terms: one that charges for spatial gradients in the field, and a second of the type introduced by Skyrme² in his soliton model of baryons, which prevents these gradients from concentrating into very small regions. The field can be represented by arrows of a fixed length that are free to point in any direction. If they are placed on a sphere, all pointing outwards, and then moved onto a plane by stereographic projection, the configuration generated is a slice through a string. The arrows rotate from 'up' outside the core to 'down' in the centre, and in between there is a winding of the horizontal component of the arrow around the core, which prevents this configuration from being smoothed out.

But simple loops of string will generally contract and decay. By introducing suitable twists and linkings, Faddeev and Niemi have produced two types of stable knot, called unknots and trefoils. Numerically solving the equations of motion, they have demonstrat-

ed that these finite loops of string do not contract or dissipate, and are thus lump solitons.

Now that these knot configurations have been constructed, their properties can be investigated in detail. One aspect of particular interest is their stability to perturbations, and the related question of how often such objects would form from random configura-

tions of string. A cosmological string model, for example, would be invalidated if it predicted copious knot production. □

Warren Perkins is in the Department of Physics, University of Wales at Swansea, Singleton Park, Swansea SA2 8PP, UK.

1. Faddeev, L. & Niemi, A. J. *Nature* **387**, 58–59 (1997).

2. Skyrme, T. H. R. *Proc. R. Soc. Lond. A* **260**, 127 (1961).

Protein transport

Greasing the Golgi budding machine

Thomas F. J. Martin

The intracellular compartment that was discovered by Camillo Golgi 99 years ago is now known to be the central station for sorting membrane and secretory proteins to their proper destinations in cells. Proteins are carried to the cell surface as cargo in transport vesicles, which form by budding of the membrane on one face of the polarized Golgi complex. A fundamental mystery of cellular protein trafficking concerns the budding process — how does it generate vesicles that are enriched for the cargo to be transported, but depleted for resident constituents, which remain in the Golgi? Spatially localized reactions that modify the phospholipid composition of the Golgi membrane could be important in budding, and our understanding of these reactions is brought to a new level by the incisive genetic and biochemical studies of Kearns *et al.*¹ on page 101 of this issue. Their work on the *Saccharomyces cerevisiae* phosphatidylinositol transfer protein Sec14p points to a key role for diacylglycerol in the process of budding (Fig. 1).

Like other cellular membranes, the Golgi membrane consists of a bilayer of luminal and cytoplasmic leaflets of orientated phospholipids (Fig. 2, overleaf). At the budding site, the surface of the membrane is coated with proteins that are recruited from the cytosol. One model for bud generation suggests that these coat proteins polymerize, and drive deformation of the membrane to form a bud that is then nipped off to generate a transport vesicle². An alternative view suggests that the phospholipid bilayer is an active, rather than a passive, participant in the dynamic changes that underlie budding³. In artificial membranes, bilayer curvature and vesicle formation can be driven by small changes in phospholipid composition⁴, resulting in lateral phase-separation into lipid microdomains or asymmetry in the monolayer leaflets. So physical changes in the membrane, caused by regulation of phospholipid metabolism, could be the primary basis for membrane dynamics.

Dramatic progress in understanding membrane and protein trafficking in eukaryotic cells has come from studies in *S. cerevisi-*

*ae*⁵. Dysfunction of the *SEC14* gene enlarges the normally inconspicuous Golgi complex by arresting the membrane traffic through the Golgi at a late stage. In 1990, Bankaitis *et al.*⁶ reported that the *SEC14* locus encodes the main yeast phosphatidylinositol transfer protein (PITP), which mediates the *in vitro* transfer of two water-insoluble phospholipids — phosphatidylinositol (PtdIns) and phosphatidylcholine (PC) — between membranes. The discovery that Sec14p is a PITP provided the first tangible link between protein transport and the phospholipid composition of the membrane compartment that is involved in the transport.

Bankaitis and colleagues identified seven genes that bypass the requirement for Sec14p when mutated. Three of these bypass suppressors are loss-of-function alleles of genes that encode the three enzymes in a pathway for the production of PC. Subsequent work indicated that PITP, in its PC-bound form, inhibits PC synthesis by regulating a rate-limiting enzyme in the pathway, suggesting

that unrestrained synthesis of PC is deleterious to the secretory competence of the Golgi⁷.

The PtdIns-bound form of Sec14p has a distinct, but complementary, function to that of PC-bound Sec14p in maintaining the activity of the Golgi. Although the mechanism is unknown, the role of PtdIns-bound Sec14p can be supplied by the structurally unrelated mammalian PITPs⁷. Mammalian PITPs are also required for membrane trafficking events and, specifically, for the formation of secretory vesicles from the Golgi⁸. Recent work indicates that the mammalian PITPs provide substrate PtdIns to lipid kinases, to form PtdIns-4-phosphate or PtdIns-3-phosphate, which are essential for signal transduction, lysosomal sorting and exocytosis⁹.

Kearns *et al.*¹ have analysed the *sec14*-bypass suppressor, *SAC1*. They provide insights into the PtdIns side of Sec14p's dualism, and further clarify its function as a PC sensor. The *SAC1* gene encodes a novel transmembrane protein, which possesses a domain that is related to inositol-phospholipid-metabolizing enzymes. Yeast strains harbouring *sec14*-bypass suppressor alleles of *sac1* show dramatic increases in the Golgi-associated synthesis of diacylglycerol. This is due to increased production of phosphoinositol-containing sphingolipids, where PtdIns serves as the inositol donor and diacylglycerol is a product. The authors infer that the increased production of diacylglycerol by *sac1* mutants is essential for restoration of Golgi secretory competence, because the *sac1* bypass is compromised and the mutant *sec14* phenotype is exacerbated by overexpression of diacylglycerol kinase. A unifying hypothesis emerges in which PITP functions by dual mechanisms — the PC-bound PITP moderates the synthesis of

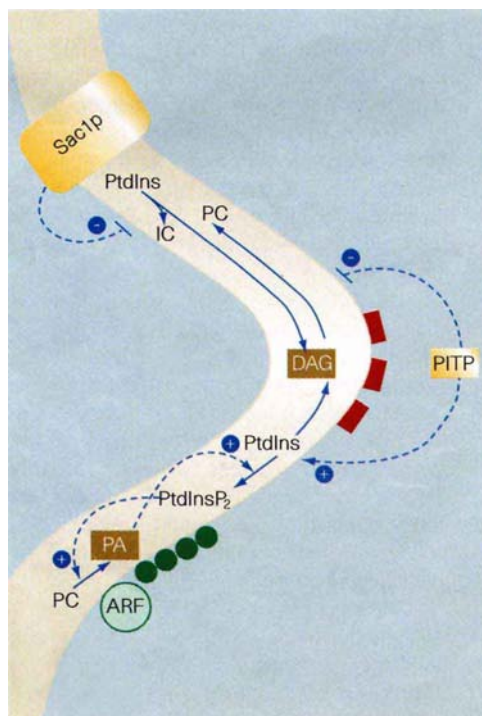


Figure 1 The formation of a vesicle from the Golgi membrane by budding is depicted as being due to localized changes in phospholipid composition that result from the generation of diacylglycerol (DAG) and phosphatidic acid (PA). Generation of DAG requires PITP (yeast Sec14p), which inhibits consumption of DAG by a phosphatidylcholine (PC)-synthesis pathway, and promotes DAG formation from phosphatidylinositol (PtdIns) metabolism. Kearns *et al.*¹ studied a mutant Sac1p that allows the PITP requirement to be bypassed by promoting the formation of DAG through increased conversion of PtdIns to inositol ceramides (IC). DAG may alter membrane curvature and recruit DAG-binding proteins (red). ARF (ADP ribosylation factor), which is needed for recruitment of coat protein (green) during budding, activates phospholipase D, which catalyses the formation of PA from PC. PA, along with the increased levels of PtdInsP₂ that may form, could function to alter membrane curvature and recruit coat proteins.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Figure 2 Budding star — electron micrograph of a circular Golgi body in a rabbit epididymis.

PC and, therefore, consumption of diacylglycerol, whereas the PtdIns-bound P1TP promotes the turnover of PtdIns, presumably to increase production of diacylglycerol. This dualism would be devoted to preserving the levels of diacylglycerol that would be needed to maintain Golgi secretory competence.

Why might budding in the Golgi require diacylglycerol? Proteins that bind diacylglycerol, for example protein kinase C, have been implicated in regulating the formation of Golgi buds in mammalian cells. Characterization of *sec14* suppressors that operate downstream of Sec14p may identify other effectors of diacylglycerol. In addition, because diacylglycerol is a non-bilayer-forming lipid, dramatic changes in the physical properties of the bilayer, including increased curvature, could initiate membrane budding¹⁰.

Other processes that modify the lipid composition of the membrane have also been linked to Golgi budding. ADP ribosylation factor (ARF) — a small G protein that is needed for recruitment of coat proteins to the Golgi — is also an activator of phospholipase D, which catalyses the hydrolysis of PC to phosphatidic acid (PA). PA activates enzymes that synthesize phosphatidylinositol-4,5-bisphosphate (PtdInsP₂), which in turn is required for activation of phospholipase D (ref. 9). Ktistakis *et al.*³ have shown that in budding, the need for ARF can be bypassed by providing phospholipase D, suggesting that coat recruitment may be driven by ARF-mediated regulation of phospholipid metabolism. The generation of membrane microdomains, enriched in

negatively charged lipids such as PA and PtdInsP₂, could promote membrane curvature for budding and serve as sites for coat-protein binding^{3,9}.

Lateral diffusion and metabolic conver-

sion may mean that the membrane domains of diacylglycerol and PA do not persist long enough to cause membrane budding. Protein binding can stabilize phospholipid domains¹¹, so coat proteins that bind acidic phospholipids and diacylglycerol-binding proteins such as protein kinase C could transform nascent buds by stabilizing them. Interactions between distinct lipids, coat proteins and integral membrane proteins to propel bud formation would result in vesicles with unique lipid and protein compositions. The new work on the importance of diacylglycerol and PA will undoubtedly spur investigations that could resolve the mysteries of Golgi budding by the end of the Golgi centenary. □

Thomas F. J. Martin is in the Department of Biochemistry, University of Wisconsin, 420 Henry Mall, Madison, Wisconsin 53706, USA.

1. Kearns, B. G. *et al.* *Nature* **387**, 101–105 (1997).
2. Rothman, J. E. & Wieland, E. T. *Science* **272**, 227–234 (1996).
3. Ktistakis, N. T., Brown, A. H., Waters, M. G., Sternweis, P. C. & Roth, M. G. *J. Cell Biol.* **134**, 295–306 (1996).
4. Sheetz, M. P. & Singer, S. J. *Proc. Natl Acad. Sci. USA* **71**, 4457–4461 (1974).
5. Schekman, R. & Orci, L. *Science* **271**, 1526–1533 (1996).
6. Bankaitis, V. A., Aitken, J. R., Cleves, A. E. & Dowhan, W. *Nature* **347**, 561–562 (1990).
7. Alb, J. G., Kearns, M. A. & Bankaitis, V. A. *Curr. Opin. Cell Biol.* **8**, 534–541 (1996).
8. Ohashi, M. *et al.* *Nature* **377**, 544–547 (1995).
9. DeCamilli, P., Emr, S. D., McPherson, P. S. & Novick, P. *Science* **271**, 1533–1539 (1996).
10. Lipowsky, R. *Nature* **349**, 475–481 (1991).
11. Glaser, M. *et al.* *J. Biol. Chem.* **271**, 26187–26193 (1996).

X-ray astronomy

Violent hiccup from a newborn star

Stephen L. Skinner

Almost all types of star emit X-rays^{1–3}. Our most detailed knowledge of this phenomenon comes from high-resolution observations of the Sun, which show that most of its X-ray emission comes from multi-million-degree plasma trapped in closed magnetic structures in the tenuous corona⁴. The Sun is a rather inefficient X-ray emitter, with only about one-millionth of its total radiative energy output emerging as X-rays. But the observations reported by Grosso *et al.* on page 56 of this issue⁵ suggest that, during its formative stage roughly 4.6 billion years ago, the young Sun may have been a much more powerful X-ray source than it is today. Intense X-ray outbursts from the proto-Sun would have subjected the inner nebula to rapid heating, probably accompanied by particle bombardment and shock waves induced by mass ejections. Such cataclysmic events may have influenced the evolution of the inner Solar System.

Using instruments on board the orbiting *Röntgensatellit* (Rosat) observatory, Grosso *et al.* detected an X-ray outburst of unprecedented intensity from a young star named YLW15, located in the Rho Ophiuchi

star-forming region some 520 light years away⁶. The estimated age of YLW15 is on the order of 100,000 years, making it a youngster on the cosmic timescale where ages are commonly measured in billions of years. The infant star, or 'protostar', is probably still in its accretion phase, and is shrouded in prenatal dust, making it invisible to optical telescopes. But its presence is known from infrared images⁶, by virtue of the heat liberated as the enveloped star undergoes gravitational collapse. Now, higher-energy X-rays, which can penetrate the dust barrier, have also been detected in a brief burst or flare whose duration was probably less than one day. Had YLW15 not undergone such an X-ray hiccup, it would have escaped detection by Rosat.

When compared with previous X-ray studies of star-forming regions, the results of Grosso *et al.* stand out for two reasons. First, the identification of YLW15 as the source of the X-ray outburst is fairly certain, owing to the accurate X-ray position obtained using the Rosat High-Resolution Imager. Previous claims of X-ray detections of protostars based on lower-resolution images of crowded star-forming cores have always been plagued by

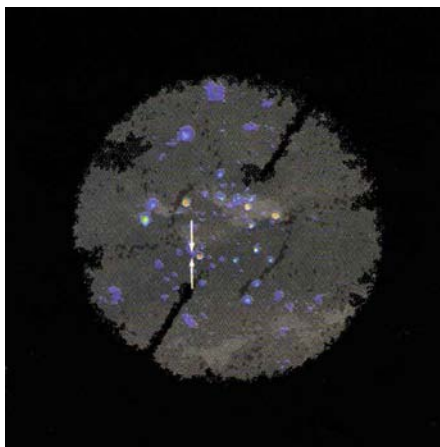


Figure 1 X-ray image of the Rho Ophiuchi star-forming region, obtained with the Rosat Position Sensitive Proportional Counter (PSPC) in 1991. The angular diameter of the image is about 2 degrees, so the region is about 18 light years across. Most of these sources are young, optically visible stars; YLW15 (arrowed), which has produced an unprecedentedly brilliant X-ray flare, is even younger, and still embedded in a dusty nebula.

ambiguities in source identification. Second, the intrinsic X-ray luminosity or power of the outburst is much higher than has been seen from anything in our Galaxy other than compact objects (neutron stars, white dwarfs and black-hole candidates). After correcting for the effects of dust absorption, the inferred X-ray luminosity is about 10^{34} – 10^{36} erg s⁻¹. The spread in values is mainly due to an imprecise knowledge of the amount of absorbing dust. But even the low end of this luminosity range would mean that YLW15 is about ten million times brighter in X-rays than the Sun, and about ten times brighter than the most luminous Galactic stellar X-ray sources, compact objects excluded.

Could the same physical processes that produce X-ray flares on the Sun (and by inference on other Sun-like stars) be responsible for the huge X-ray output of YLW15? Grosso *et al.* think not, arguing that solar mechanisms are too inefficient and that the inferred size of the X-ray-emitting region of YLW15 is much larger than the protostar itself. Although the energy release mechanism in solar X-ray flares is still controversial, evidence is gathering^{7,8} that it is a process called magnetic reconnection. In this picture the magnetic field rapidly drops to a lower energy state as the result of a localized rearrangement of magnetic field lines (for further discussion see ref. 9).

Some of the field energy released in this process goes into heating the plasma near the reconnection site, which then cools rapidly by emitting thermal X-rays. In the Sun, the magnetic structures that reconnect are confined to the vicinity of the solar surface. Grosso *et al.* argue that, in contrast, the X-ray emission from YLW15 may originate when

magnetic field lines anchored to the protostar reconnect with an external structure such as a circumstellar disk or perhaps a protostellar companion. That would increase the length scale of the participating magnetic structures to tens of stellar radii or more, as deduced from models of the X-ray outburst.

Several new questions are raised by this brilliant flare. If the energy release does involve magnetic processes, then how are the magnetic fields produced in such a young object? Are they internally generated, or are we perhaps witnessing the venting of primordial magnetic energy that the protostar inherited from its parent molecular cloud? And what sort of external structure might the protostellar magnetic field be interacting with? Any answer must be considered speculative, but one of the most enticing possibilities is that we are seeing a magnetic interaction between the two components of a close protostellar binary system — the birth of twins. Some indirect evidence in support of this idea comes from the fact that YLW15 is one of the strongest radio sources in Rho Ophiuchi, and at least one study suggests that the frequency of binary systems in this star-forming region

may be higher among radio emitters¹⁰.

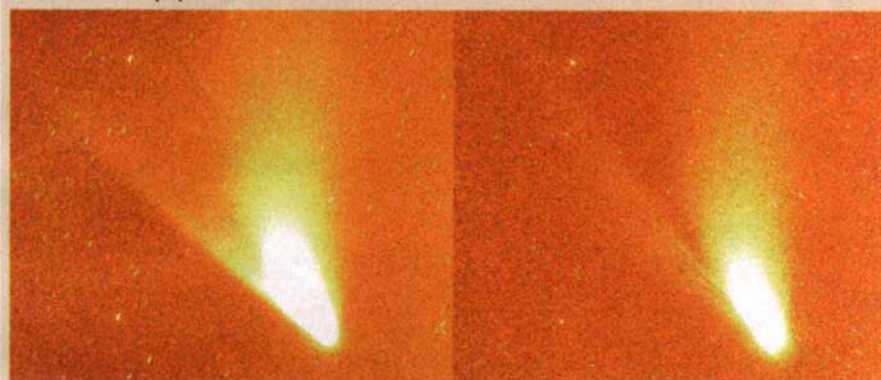
Finally, if our own Sun once experienced such convulsive releases of X-ray energy, what would have been the implications for the chemistry, thermal structure and evolution of the inner nebular disk where the terrestrial planets formed? Current models of nebula evolution¹¹ do not consider the cumulative effects of energy input from intense X-ray flares, but such refinements may now be necessary. □

Stephen L. Skinner is at JILA, University of Colorado, Boulder, Colorado 80309-0440, USA.

1. Haisch, B. & Schmitt, J. H. M. M. *Publ. Astron. Soc. Pacif.* **108**, 113–129 (1996).
2. Rosner, R., Golub, L. & Vaiana, G. S. *Annu. Rev. Astron. Astrophys.* **23**, 413–452 (1985).
3. Charles, P. A. & Seward, F. D. *Exploring the X-ray Universe* (Cambridge Univ. Press, 1995).
4. Acton, L. *et al. Science* **258**, 618–625 (1992).
5. Grosso, N. *et al. Nature* **387**, 56–58 (1997).
6. Wilking, B. A., Lada, C. J. & Young, E. T. *Astrophys. J.* **340**, 823–852 (1989).
7. Masuda, S., Kosugi, T., Hara, H., Tsuneta, S. & Ogawara, Y. *Nature* **371**, 495–497 (1994).
8. Innes, D. E., Inhester, B., Axford, W. I. & Wilhelm, K. *Nature* **386**, 811–813 (1997).
9. Klimchuk, J. *Nature* **386**, 760–761 (1997).
10. Strom, K. M., Kepner, J. & Strom, S. E. *Astrophys. J.* **438**, 813–829 (1995).
11. Cassen, P. in *Chondrules and the Protoplanetary Disk* (eds Hewins, R. H. *et al.*) 21–28 (Cambridge Univ. Press, 1996).

Comets

Hale-Bopp's atomic tail



Hale-Bopp is fading, but on its way out it has sprung a surprise. It has got a third tail, a tail made of sodium — something never before seen.

A team using the Isaac Newton Telescope on La Palma looked for emission from sodium atoms using a spectral filter. With the filter on, they took the picture on the left: the sodium tail is shown, clearly much straighter and narrower than the wide, bright, curving dust tail to its right.

Without the filter (right) the ion tail is visible

as a faint, fan-like structure. The charged ion tail is driven by the solar wind, but that should have no effect on neutral sodium atoms. So what forms the sodium tail? It may be that, like the dust in the dust tail, the sodium atoms are accelerated by light pressure. But instead of reflecting the Sun's light like the dust, they absorb it, and so are accelerated by fluorescence (G. Cremonese, H. Rauer & A. Fitzsimmons *IAU Circ. No.* 6634; 1997).

Unfortunately, the new tail is too faint to see with

the naked eye, but its colour should be familiar. It is the same yellowish colour — a pair of spectral lines at 589 nanometres — produced by low-pressure sodium street lamps, which have become widespread in many countries as a source of ugly lighting. But sodium lights are dimmer than their tungsten or mercury-vapour predecessors, so contribute less to light pollution. Those of us who live in large cities owe the visibility of Hale-Bopp, in some slight part, to sodium.

Stephen Battersby

The ancestry of segmentation

E. M. De Robertis

Until recently, most evolutionary biologists would have agreed that the segments of an arthropod, such as an insect, and of the backbone of a vertebrate were independent ('convergent') solutions to a functional need for building bilaterally symmetrical body plans using modular units. The work reported by Linda Holland and colleagues in the May issue of *Development*¹ is bound to change that view.

Holland and colleagues find that a homologue of the *engrailed* gene is expressed in the posterior half of each of the first eight segments (somites) of amphioxus (Fig. 1). In *Drosophila*, the *engrailed* gene encodes a DNA-binding protein that specifies the posterior compartment of each segment. Amphioxus has a notochord but not vertebrae, and is very useful in evolutionary studies because it appears to have retained many of the characteristics of the archetypal chordate ancestor that gave rise to the vertebrates². The fact that *engrailed* is expressed in both *Drosophila* and chordate metameres¹ tells us that segmentation was present in the common ancestor from which the insect and chordate lineages diverged 500 million years ago,



Figure 1 Expression of *engrailed* mRNA in the posterior portion of each of the five somites that have been formed in a 12-hour amphioxus embryo. The sixth stripe is in the somite that is about to be formed in the posterior (to the right). (Photo courtesy of Linda Holland, from ref. 1.)

the Urbilateria (Ur, primaevus; Bilateria, bilateral animal)³. The report on amphioxus *engrailed* follows soon after the discovery of *her1*, a homologue of the hairy DNA-binding protein that is expressed in every other forming somite in the zebrafish⁴. In *Drosophila*, *hairy* is a 'pair-rule' gene required for the formation of alternating segments. The pair-rule expression of *her1* in zebrafish led Kimmel to propose that Urbilateria was segmented⁵.

The Nobel prize-winning screen for *Drosophila* embryonic lethal mutations⁶ provided researchers with many pair-rule and segment-polarity genes, but their homo-

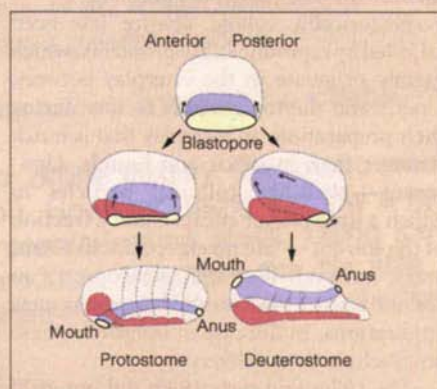
logues were not implicated in the formation of vertebrate somites. In fact, many homologues of *Drosophila* segment-polarity genes (such as *engrailed*, *wingless*, *armadillo*, *hedgehog* and *patched*) have been found to have other functions in vertebrates, such as the formation of the midbrain-hindbrain border and dorsoventral patterning of the neural tube. The new work¹ will undoubtedly stimulate a second, hard look for expression of these homologues in forming somites in regions such as the tailbud of the vertebrate embryo.

In zebrafish, chick and mouse, *engrailed* homologues are expressed in subsets of cells in the somites^{7,8}, but only after somites have been formed. This is different from amphioxus *engrailed* or zebrafish *her1*, which form bands of expression before morphological signs of segmentation are detectable^{1,4}. There are three *engrailed* genes in zebrafish and only one in amphioxus^{1,7}. Although one could argue that an as yet undiscovered *engrailed* homologue might explain the failure to detect early expression, it is more likely that the differences from amphioxus are due to the mode of somite formation. Amphioxus has 50 or more somites but the *engrailed* bands are observed only in the first eight, which are formed by a primitive mechanism by which epithelial outpockets are pinched off the primitive gut ('enterocoely'), whereas more posterior somites are formed from mesenchymal cell blocks which are then subdivided into seg-

Evolving protostomes and deuterostomes

Zoologists classify bilateral animals that possess a coelomic cavity into two fundamental groups: the protostomes – such as arthropods, annelids and molluscs; and the deuterostomes – such as echinoderms and chordates (chordates include ascidians, amphioxus and vertebrates). In protostomes, the mouth is formed at or close to the initial site (proto, first; stomo, mouth) of the blastopore, the site at which the endomesoderm, shown here in mauve, involutes into the interior of the embryo. In many protostomes the anus is formed at the blastopore as well. In deuterostomes, the mouth is formed secondarily (deutero, second) by a perforation of the ectoderm, and the anus is formed at or close to the site of the original blastopore. In addition, most protostomes have a ventral nerve cord which is traversed by the mouth and connects to a supraesophageal ganglion (brain) in the anterior. In most deuterostomes the central nervous system (CNS, in red) is dorsal and is not traversed by the gut. There are other differences, but these suffice for the present discussion.

Developmental studies suggest that protostomes and deuterostomes had a



common ancestor that was complex and segmented, raising the question of how such different body plans evolved. We of course do not know how the adult Urbilateria looked; it could have had either an open gut/blastopore (not shown) or it could have resembled an adult protostome or deuterostome (bottom diagrams). However, by modifying the blastopore during gastrulation it is possible to envisage how the transition could have happened, provided one assumes that the CNS is formed near the blastopore⁹.

As shown in the figure, in protostomes such as annelids the slit-

like ventral blastopore gives rise to the mouth and anus by closing along its central portion. The fusion of the lateral blastopore lips results in the dorsoventral inversion of the CNS, generating posteriorly a nerve cord ventral to the gut, and anteriorly a supraesophageal ganglion (bottom left). In deuterostomes the anterior part of the blastopore, corresponding to the mouth, does not form at all and endomesodermal involution takes place from the posterior, eventually leading to formation of the anus. The CNS is induced in nearby ectoderm by proteins secreted by the invaginating endomesoderm, and by the end of gastrulation the deuterostome mouth is perforated secondarily, with the gut and CNS remaining in different sides of the animal throughout their length. The deuterostome depicted here is a frog tadpole, which if inverted would adopt its normal dorsoventral position. Thus, although the Urbilateria ancestor was complex in its adult form (presumably having segments, heart, eyes and appendages), the potential to give rise to widely divergent body plans resided in its mechanisms of embryonic development. E.M.DeR.

ments. The latter mode, formation of somites from a block of mesenchyme, is observed throughout the vertebrates⁹ and may explain the lack of early *engrailed* expression¹.

Even before the recent gene-marker studies^{1,4}, another similarity should have alerted us to the possibility of a common origin of insect and vertebrate segmentation. In 1855 Remak described the 'resegmentation' process by which vertebrae are formed¹⁰. In mammalian somites a subset of cells (the sclerotome) forms the vertebrae, whereas the rest gives rise to muscle and dermis (dermomyotome). Each mature vertebra is formed by the posterior half of one sclerotome which fuses to the anterior half of the next one. The end result is a phase shift of the vertebra with respect to the muscle, so that the segmental muscles can span, and move, adjoining vertebrae.

Although sometimes disputed⁹, vertebral resegmentation has recently been confirmed by cell-lineage studies using the chick-quail system (L.-M. Bourcheix and N. Le Douarin, personal communication). In an analogous fashion, in *Drosophila* the initial segmentation unit is the parasegment which subsequently subdivides and forms the posterior compartment of one segment and the anterior compartment of the next definitive segment. Pair-rule genes such as *hairy* are required to regulate this two-step segmentation process¹¹. It seems improbable that such a complicated way of making individual metameres would have arisen independently twice in evolution.

So three lines of evidence — *engrailed* in amphioxus, a pair-rule *hairy* homologue in zebrafish, and resegmentation in *Drosophila* and vertebrates — support the notion that segmentation in insects and chordates is homologous (that is, derived by descent from a common ancestor). In addition to segmentation, studies with conserved developmental control genes (indicated in parentheses; see ref. 3 for details) suggest that Urbilateria also had the following characteristics: antero-posterior polarity (*Hox* gene complexes), dorsoventral patterning (*sog/chd* and *dpp/BMP-4*), a primitive photoreceptor (*Pax6/eyeless*), a contractile blood vessel or heart (*Tinman/Nkx 2.5* and *DMEF2*), and perhaps a humble appendage or antenna-like outgrowth (*fringe*, *serrate* and other genes¹²).

The reader might rightly ask how it was possible to evolve the enormous variety of bilateral animals that surround us in our daily lives from such a complex ancestor. The answer lies not in the adult forms but in the embryo, in particular in the modifications of embryonic development that gave rise to the protostomes and the deuterostomes¹³. This divergence into two main subdivisions of animals (see box on page 25) took place early in evolutionary time, as indicated by ribosomal DNA phylogenetic trees¹⁴. The realization that all Bilateria are derived from a complex ancestor represents a major change in

evolutionary thinking, suggesting that the constraints imposed by the previous history of species played a greater role in the outcome of animal evolution than anyone would have predicted until recently. □

E. M. De Robertis is in the Howard Hughes Medical Institute, University of California, Los Angeles, California 90095-1662, USA.

1. Holland, L. Z., Kene, M., Williams, N. A. & Holland, N. D. *Development* **124**, 1723-1732 (1997).
2. Gee, H. *Before the Backbone: Views on the Origin of the Vertebrates* (Chapman & Hall, London, 1996).
3. De Robertis, E. M. & Sasai, Y. *Nature* **380**, 37-40 (1996).

4. Müller, M., von Weizsäcker, E. & Campos-Ortega, J. A. *Development* **122**, 2071-2078 (1996).
5. Kimmel, C. B. *Trends Genet.* **12**, 320-331 (1996).
6. Nüsslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795-801 (1980).
7. Ekker, M., Wegner, J., Akimenko, M. A. & Westerfield, M. *Development* **115**, 1001-1010 (1992).
8. Davis, C. A., Holmyard, D. P., Millen, K. J. & Joyner, A. L. *Development* **111**, 287-298 (1991).
9. Keynes, R. J. & Stern, C. S. *Development* **103**, 413-429 (1988).
10. Remak, R. *Untersuchungen über die Entwicklung der Wirbeltiere* (Reimer, Berlin, 1855).
11. Lawrence, P. A. *Development* **104**, Suppl. 61-65 (1988).
12. Laufen, E. et al. *Nature* **386**, 366-373 (1997).
13. Arendt, D. & Nübler-Jung, K. *Mech. Dev.* **61**, 7-21 (1997).
14. Aguinaldo, A. M. A. et al. *Nature* (in the press).

Colloid chemistry

Liposomes within liposomes

Danilo D. Lasic

Living nature performs many of its tasks using the self-assembling organization of polar lipid molecules. Each eukaryotic cell is a complex multicompartiment colloidal system. The paper by Walker *et al.*, on page 61 of this issue¹, describes the artificial preparation of similar structures: small vesicles are aggregated and then wrapped by a larger membrane, using the property that some lipids have of being able to change from an open, rolled-up bilayer into a large vesicle when electrostatic conditions change in the solution². This is an example of the rational construction of complex liposomes, using the interaction between surface-attached ligands and soluble receptors, and a new method for wrapping the aggregates.

Historically, colloid science has been crippled by reproducibility problems, which mainly originate in the interplay between kinetic and thermodynamic factors during their preparation, an interplay that is much stronger than in solids and liquids. Liposomes — shell-like colloidal particles in which a lipid bilayer encapsulates a fraction of the solvent — are no exception. In recent years, however, the rapid development in colloid science has resulted in many new applications. In the case of liposomes these are chiefly in drug delivery.

The 1970s and early 1980s did not yield any viable products, but the fundamental research, which followed a mostly trial-and-error approach, revealed many basic characteristics of such systems, both in the test tube and in multicompartiment structures in living organisms³. The improved understanding of their stability and interaction characteristics paved the way for a scientifically based rational approach in liposome design. Now we can understand liposomes and their properties via several physically measurable quantities, such as the bending and stretching elasticity of the lipid bilayer, and attractive and repulsive interactions on a molecular and a colloid level⁴.

The development of classical liposomes and other lipid-based drug carriers culmi-

nated in three formulations of a very insoluble drug (amphotericin B, for fungal infections) embedded in the bilayer upon specific interaction with a particular lipid in well-defined conditions. But it was only the synergistic action of polymers associated with liposomes that revived liposomal drug delivery. A polymer coating results in sterically stabilized liposomes, which have proved to be very effective in cancer chemotherapy. Additionally, the stability and interaction characteristics of these liposomes are well understood. Recently, for instance, it was shown that by breaking the symmetry of polymer distribution in each leaflet of the bilayer, a spontaneous curvature can be generated, and vesicles of extremely narrow size distribution have been formed simply by increasing the salinity of special micellar mixtures⁵.

Walker *et al.*¹ have now achieved self-assembly of many-compartment liposomes, with relatively high encapsulation efficiency, by attaching lipid rolls to vesicle aggregates by biotin-streptavidin linkages⁶. Why is this important? Although the main goal was to decrease the leakage of encapsulated agents, their work is also an elegant way to encapsulate larger particles such as proteins or nucleic acids into liposomes — a major problem for many applications. Moreover, vesicles with different functionalities can be combined in this structure: the larger liposome might deliver a load of highly active, smaller liposomes, perhaps containing highly toxic drugs, to a specific site, thus sparing other tissues from any side effects. Such a system could be important in the treatment of cancer, but first the liposomes will have to be made smaller than 250 nm (so far, the smallest are about 300 nm across) and coated in polymers to increase their survival time in the blood.

Other liposomes have been successfully coated in this way. The obvious next step is to attach various ligands to them — ligands can carry out various functions, such as specifically targeting appropriate receptors *in vivo*, and thus choosing which membranes to fuse with. Although natural barriers and defence

mechanisms still present enormous problems, some useful applications, at first perhaps in diagnostics, may follow within the next few years. To further control the interaction characteristics of liposomes, such as interactions with cells or response to drug leakage, liposomes with time-dependent characteristics are being developed. These include shedding of the polymer coat or a phase transition to a non-bilayered phase upon some chemical or physical change.

Another new direction is the development of complexes of cationic liposomes with DNA plasmids for gene therapy, a treatment based on selectively switching various genes on or off. Genetic defects have been identified and appropriate DNA plasmids prepared; the problem is their effective and safe delivery into appropriate cells *in vivo*. But cationic lipids were found to be effective carriers⁷.

After a decade of trying to solve the delivery problem by a mostly trial-and-error approach, the first thorough physico-chemical characterizations of these complexes have been published. X-ray studies revealed that there is much more order in these structures

than previously thought^{8,9}. Lamellar structures were shown to contain two-dimensionally condensed and ordered DNA sandwiched between cationic bilayers, and correlations were found between transfection activity and DNA condensation¹⁰.

Liposome-based drug delivery has come of age in the mid-1990s, and more improvements are to come in our basic understanding of colloidal lipid systems, and in the design and application of intelligent colloids. □

Danilo D. Lasic is at Liposome Consultations, 7512 Birkdale Drive, Newark, California 94560, USA (e-mail: dan.lasic@roche.com).

1. Walker, S. A., Kennedy, M. T. & Zasadzinski, J. A. *Nature* **387**, 61–64 (1997).
2. Papahadjopoulos, D., Vail, W., Jacobson, K. & Poste, G. *Biochim. Biophys. Acta* **394**, 483–491 (1975).
3. Lipowsky, R. & Sackmann, E. (eds) *Handbook of Biomembranes* (North Holland, Dordrecht, 1995).
4. Lipowsky, R. *Nature* **349**, 475–481 (1991).
5. Joannic, R., Auvray, L. & Lasic, D. D. *Phys. Rev. Lett.* (in the press).
6. Chiruvolu, S. *et al. Science* **264**, 1753–1756 (1994).
7. Felgner, P. L. *et al. Proc. Natl Acad. Sci. USA* **84**, 7413–7417 (1987).
8. Lasic, D. D. *et al. J. Am. Chem. Soc.* **119**, 832–833 (1997).
9. Rädler, J. O., Koltover, L., Salditt, T. & Safinya, C. R. *Science* **275**, 810–814 (1997).
10. Lasic, D. D. *Liposomes in Gene Delivery* (CRC Press, Boca Raton, FL, 1997).

Obesity

Some heat but not enough light

Jules Hirsch

The increase in the prevalence of obesity seems to be unaffected by the warnings of public-health officials, nutritionists and physicians. Obviously, there is no way to accumulate excess adipose tissue without an inequality between the intake and expenditure of calories. Scientists have concentrated mainly on calorie intake — neural mechanisms of intake control and satiety, diet, appetite-regulating drugs and even behaviour modification. Most recently, it has been shown that adipocytes produce a substance called leptin, which is important in controlling food intake, at least in mutant rodents. Now, on pages 90 and 94 of this issue, reports by Enerbäck *et al.*¹ and Thomas and Palmiter² turn attention to the expenditure of calories.

The expenditure side of the equation is generally thought of as physical activity, but this is not the whole story. Enerbäck *et al.*¹ have studied a potential means of body-mass regulation, formerly known as 'luxus consumption'. More than 85 years ago, German clinicians wondered whether we might have some mechanism for burning away our dietary indulgences, thereby maintaining a fixed adipose composition. Those who are defective burners become obese, and this could explain the frequent observation that different people seem to eat more or less than others, but only some become fat.

In 1931, Wiley and Newburgh³ carefully measured weight changes in humans on various diets, and they estimated heat output

in both obese and non-obese people. They could not show that there was any excess heat or caloric wastage in people on high-calorie diets, and they concluded that obese people do not have a deficiency of caloric expenditure. From these authoritative studies came the dictum that all obesity comes about by excess intake of calories.

But new findings led to a re-examination of calorie expenditure. The existence of discrete brown-fat deposits had been known for some time. These are most easily discerned in hibernating animals, but they are also present in the young of many species — even in newborn humans. Until about 1960, the function of these mitochondrion-filled (and hence brown) collections of multi-locular fat cells was unknown. Then, brown adipose tissue was found to be a remarkable heat producer. Mitochondria have an internal structure that is uniquely suited to the passage of electrons from metabolic substrates to oxygen (Fig. 1). As electrons move across the internal mitochondrial membrane, there are at least three sites where energy is stored by the development of a proton gradient, and this gradient is essential for coupling combustion to the synthesis of ATP (oxidative phosphorylation).

In brown adipose tissue, uncoupling occurs, so heat rather than ATP is produced — this occurs during hibernation. Can diet do the same? Rothwell and Stock⁴ showed that rats that are overfed a cafeteria diet (presumably similar to what many of us eat) will



100 YEARS AGO

Four common species of the family Dermestidae have become vegetarians... Cyclopædias and textbooks inform us that the members of this well-known family feed upon dried animal substances. The depredations of certain species on leather, hides, and dried meats; of others on carpets, furs, and woollen goods; and of still others on dried insects, and other "objects of natural history" are, unfortunately, too well known to require further comment. Within recent years, however, several household dermestids have been suspected of living in the larval condition upon vegetable substances, and the charge of having vegetarian proclivities has now been brought home to four species, viz. *Attagenus piceus*, or black carpet beetle; *Trogoderma tarsale*, a bad cabinet pest; *Trogoderma sternale*, and *Anthrenus verbasci*. The first-named is found guilty of feeding upon flour and meal; the second revels in fiery-red pepper; and the third species appears to be able to thrive on such laxative substances as castor-beans and flax-seed. The change from a natural animal-feeding habit to a vegetable one is attributed to altered environment.

From *Nature* 29 April 1897.

50 YEARS AGO

The recent note "Aspergillin: a Name Misapplied to Several Different Antibiotics", in which Tobie deprecates the practice of assigning names to antibiotics without due consultation of the literature, prompts me to point out, for its historical interest, that even penicillin has not escaped such a duplication in naming. I refer to the use of the name 'penicillin' by Palei and Osuicheva in 1936 to designate a thermostable substance isolated from *Penicillium luteum purpurogenum*, which was unfavourable to the production of citric acid by *Aspergillus niger*. This was seven years after the first use of the term by Fleming in his now famous paper. The oversight is probably due in part to the vagaries of indexing. The name 'penicillin' did not appear in *Chemisches Zentralblatt* until 1934, or in the British abstracts until 1933; it can, however, be found in the 1929 *Chemical Abstracts* and the *Biological Abstracts* for 1930. This serves as another warning against too much reliance on a single abstracting source.

From *Nature* 3 May 1947.

gain weight and increase the production of heat from brown adipose tissue. Because adult animals tend to lose brown fat, its function in fully grown humans has been debated. Now Enerbäck *et al.*¹ present a new look using molecular genetics.

The protein that is responsible for uncoupling combustion and ATP synthesis in brown adipose tissue is the mitochondrial uncoupling protein (UCP1; formerly UCP). The structure of the *Ucp1* gene is known, so a DNA construct was developed containing the regulatory elements of this gene, along with the coding sequence for the A chain of diphtheria toxin. Transgenic mice containing this so-called 'toxigene' had greatly reduced amounts of UCP in their brown adipose tissue⁵. One transgenic strain developed obesity by increasing food intake, rather than by a reduction of energy expenditure alone. Enerbäck *et al.* have now gone one step further, by producing a *Ucp1* knockout mouse. These mice show the expected inability to keep warm in the cold but, surprisingly, they do not become obese.

This would seem to be the end of luxury consumption, but to the rescue comes the finding that there is another UCP — the mouse *Ucp2* gene was described by Fleury *et al.*⁶ in the March issue of *Nature Genetics*. It has marked homology with a human *UCP2* gene, and it is found not only in brown fat, but in white fat as well. Moreover, expression of *Ucp2* is upregulated in white adipose tissue when the mice are fed a high-fat diet. The obvious next steps are to make knockouts of *Ucp2* and to determine whether UCP2 is normally involved in burning away extra calories, and whether it is dysfunctional in the obese.

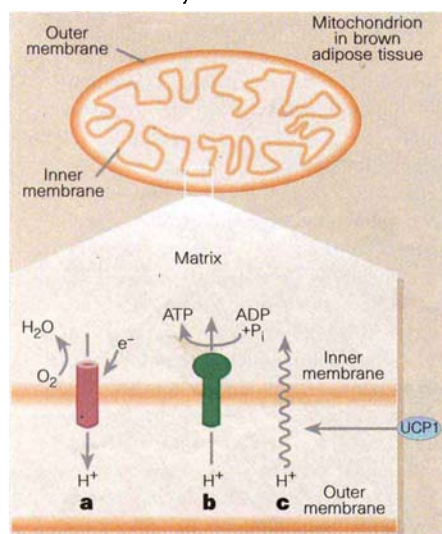


Figure 1 Oxidative phosphorylation takes place in the inner membranes of mitochondria in brown adipose tissue. **a**, An excess of protons is generated on one side of the membrane as electrons move along a 'respiratory' chain. **b**, The proton excess leads to 'coupled' oxidation, which results in the formation of ATP. **c**, The mitochondrial uncoupling protein (UCP1) creates a 'leak', dissipating the proton excess to yield heat.

In spite of these findings, the long search to find the heat producer that keeps its lucky possessors svelte is far from over. We may soon learn how UCP1 is involved in caloric exchange, but its roles in obesity treatment and pathogenesis are uncertain. Previous efforts to treat obesity by uncoupling oxidative phosphorylation using thyroid hormone, dinitrophenol or massive doses of sympathetic agonists ended in doing more mischief than good because mitochondrial uncoupling occurred at unwanted, as well as at desired, sites.

Concentrating on calorie intake or expenditure alone may also be misleading. The caloric inequalities that produce obesity are probably small and, when obesity is achieved, there may be re-equilibration to maintain it at a fixed level. This could indicate the existence of an 'equalizer' that maintains a given level of fat storage. Examples of known equalizers include the collection of elements that maintain constant blood pressure or blood glucose levels. Within the equalizer system there may be luxury consumption, which could act interdependently with all other factors. Notably, both obese and non-obese people at their usual body weight require roughly the same number of calories per unit of metabolic mass⁷. But when these people are experimentally raised 10 per cent above or below their usual weights, there is induced caloric inefficiency at the high weight (which tends to restore them to their initial state), and increased caloric efficiency at the lower weight.

Caloric wastage after weight increase is a form of luxury consumption. However, obese people behave exactly the same as the non-obese when they are changed from their obese weight — the equalizer system ordains what the level of fat storage will be, rather than this single element. Thomas and Palmiter² take this further and suggest that the equalizer has redundancy. They have generated knockout mice that lack the gene for dopamine β -hydroxylase (*dbh*), so they cannot synthesize noradrenaline or adrenaline. The effect of this is to paralyse the sympathetic nervous system, which is believed to be an essential regulator of energy metabolism. However, the authors find that knocking out the *dbh* gene has no effect on fat storage. So if the concept of the pathogenesis of human obesity is correct, examining all components in the regulatory system remains the only way to approach an understanding of the entire system. But, as of this moment, stick to your diet, keep up your jogging, and wait for the continuation of this fascinating story. □

Jules Hirsch is at the Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA.

1. Enerbäck, S. *et al.* *Nature* **387**, 90–94 (1997).
2. Thomas, S. A. & Palmiter, R. D. *Nature* **387**, 94–97 (1997).
3. Wiley, F. H. & Newburgh, L. H. *J. Clin. Invest.* **10**, 733–744 (1931).
4. Rothwell, N. J. & Stock, M. J. *Nature* **281**, 31–35 (1979).
5. Lowell, B. B. *et al.* *Nature* **366**, 740–742 (1993).
6. Fleury, C. *et al.* *Nature Genet.* **15**, 269–272 (1997).
7. Leibel, R., Rosenbaum, M. & Hirsch, J. *N. Engl. J. Med.* **332**, 621–628 (1995).

Daedalus

The kindest cut of all

Metal cutting is a brutal process. Lathes, drills and milling machines all force a cutting edge through a resistive workpiece. That edge rapidly wears from the enormous friction of the metal flowing past it under extreme pressure. Furthermore, the flow is quite unlubricated. The fiercest spray of cutting fluid cannot reach the point where the edge is parting the metal. All it can do is absorb the resulting heat.

Daedalus hopes to change all this. On a microscopic scale, he says, a cutting edge is very rough and blunt, with a highly uncertain radius. As it is driven forward, the metal shears in front of it, so that there is always a tiny gap ahead of its rounded nose. Daedalus wants to pump lubricant into that gap. Flowing back along the faces of the cutting tool, it will lubricate the twin streams of sheared metal. So he is applying modern microfabrication methods to chips, not of silicon, but of tool steel and tungsten carbide.

His new photo-etched tool chips have cutting edges of precisely defined radius. Along each edge is a series of etched holes a few micrometres across. When the chip is cemented onto its matching support, a channel in the latter reaches the holes in the chip. The resulting drill bit, lathe tool or milling cutter can be fed with lubricant through its shank. The fluid will weep through the holes and emerge along the cutting edge of the tool, into that minute gap where the metal of the workpiece is being parted. For the first time, the cutting edge itself will be lubricated.

The lubricant poses problems of its own. To leak at a useful rate from such tiny holes, it must combine the highest density with the lowest viscosity. A supercritical fluid should fill this bill — carbon dioxide, say, or nitrous oxide. An edge exuding some 70 atmospheres of an almost incompressible but readily expansible fluid should practically eliminate frictional contact at the cutting edge. A little oil dissolved in the fluid would precipitate out as it expanded, giving more conventional lubrication downstream. The strong cooling of expansion would be useful too.

The new tools will transform engineering. Lathes and milling machines will carve effortlessly through the toughest metal on a fraction of their previous power. The tools, protected by supercritical lubrication, will not wear; their products, formed by edges of optical quality, will themselves have a precision mirror finish. Metal cutting will become a delicate, quiet, contemplative form of industrial sculpture.

David Jones

A Sun compass in monarch butterflies

Each autumn, monarch butterflies (*Danaus plexippus*) migrate up to 4,000 km from breeding grounds in the eastern United States and Canada to overwintering sites in Mexico^{1,2}. We tested the ability of monarchs to orient using a Sun compass by clock-shifting³ the butterflies and observing the orientation of their subsequent flight. Clock-shifted migrants reoriented their bodies in the predicted clockwise direction whereas control groups of migrants did not. This indicates that monarch butterflies are able to orient during their transcontinental migration using a Sun compass.

In conjunction with a knowledge of the time of day, many animals are able to use the Sun's changing azimuth to orient their migratory and homeward movements⁴. Beginning with Schmidt-Koenig³, Sun compass use has been demonstrated by clock-shifting experiments. A change in the internal clock causes animals to misinterpret the position of the Sun, and hence change their direction of movement in a predictable way. In the Northern Hemisphere, for example, time-delayed migrants exhibit a clockwise change in orientation.

During September 1996, in eastern Kansas, we experimentally clock-shifted (delayed) migrating monarchs by six hours to determine whether these butterflies use a time-compensated Sun compass. We released clock-shifted subjects individually and watched them for 1–5 min. We ran behind and beneath butterflies to estimate body orientations (headings) maintained for 15 s or more, using hand-held compasses. Migrants that were kept in the laboratory but not clock-shifted served as a sham control group and naturally migrating butterflies served as 'natural' controls.

The mean heading for clock-shifted butterflies was toward the west-northwest (Fig. 1a), which is 75° from the mean south-southwest orientation of the sham controls

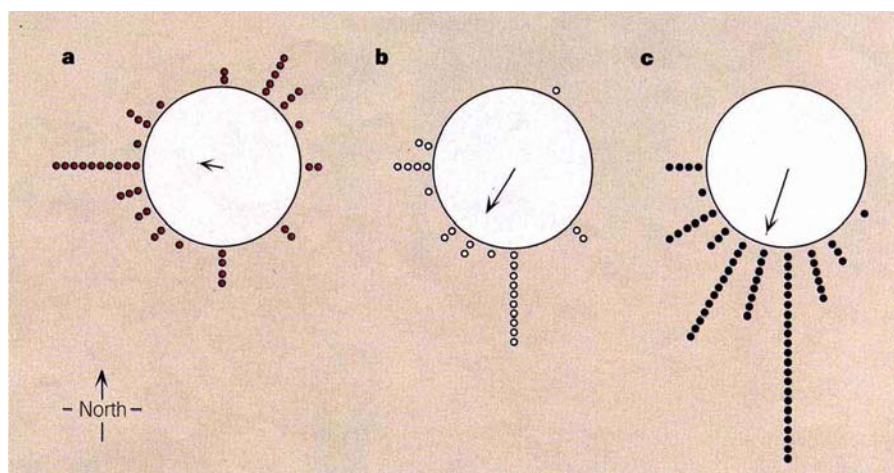


Figure 1 Mean body orientation data, giving resultant vector direction (μ) and length (r) for a, clock-shifted migrants ($\mu = 287^\circ$; $r = 0.29$; $n = 43$, each circle represents 1 migrant); b, sham controls ($\mu = 211^\circ$; $r = 0.67$; $n = 25$, each circle represents 1 control); and c, natural controls ($\mu = 200^\circ$; $r = 0.86$; $n = 204$, each circle represents 3 controls). Subjects not exhibiting directional flight for any reason (for example, spiralling upwards on thermal currents, stopping to feed) were excluded from the analyses. Of 532 migrants observed, 272 yielded stable heading data.

(Watson's F -test, $F = 13.84$, $P < 0.01$; Fig. 1b) and 85° from the mean natural control measure ($F = 49.02$, $P < 0.01$; Fig. 1c). The mean heading of the clock-shifted butterflies bears out the clockwise change in direction predicted by the clock-shift delay experienced, and is of the magnitude predicted by the 6-hour time shift. Importantly, the two control groups exhibited strikingly similar mean headings ($F = 1.93$, $P > 0.05$), indicating that there was no significant effect of captivity.

Monarch butterflies in North America use a Sun compass during their southward autumn migration, orienting their bodies using the time of day and the position of the Sun. The monarch butterfly thus joins the small group of species for which a Sun compass orientation mechanism has been demonstrated experimentally. In the absence of celestial cues on overcast days, however, monarchs still manage to orient

towards the south-southwest⁵, suggesting that they also have a non-celestial backup mechanism of orientation, such as a geomagnetic compass^{6,7}.

Sandra M. Perez

Department of Ecology and Evolutionary Biology,
University of Arizona,
Tucson, Arizona 85721, USA
e-mail: mperez@u.arizona.edu

Orley R. Taylor

Rudolf Jander

Department of Entomology,
University of Kansas,
Lawrence, Kansas 66045, USA

1. Urquhart, F. A. *Can. Entomol.* **73**, 71–72 (1941).
2. Brower, L. P. *J. Exp. Biol.* **199**, 93–103 (1996).
3. Schmidt-Koenig, K. *Naturwissenschaften* **45**, 47 (1958).
4. Wehner, R., Lehrer, M. & Harvey, W. R. (eds) *J. Exp. Biol.* **199**(1), (1996).
5. Schmidt-Koenig, K. *Behav. Process.* **4**, 73–78 (1979).
6. Baker, R. R. *Anim. Behav.* **35**, 94–101 (1987).
7. MacFadden, B. J. & Jones, D. S. in *Magnetite Biomineralization and Magnetoreception in Organisms* (eds Kirschvink, J. L., Jones, D. S. & MacFadden, B. J.) 407–415 (Plenum, New York, 1985).

Electrical recording from a glass sponge

Sponges arose very early in metazoan evolution. They do not have a nervous system, but some respond to stimulation by producing local contractions and one group, the 'glass sponges' (Hexactinellida), shows coordinated arrests of movements of the flagella, which produce the feeding current^{1,2}. We show here that these arrests are coordinated by propagated electrical impulses. This is, to our knowledge, the first

report of electrical signalling in any sponge.

Previous attempts to record electrical activity from sponges have been hampered by the delicate, porous nature of the tissues, which makes attachment or insertion of electrodes difficult. In the sponge that we studied (*Rhabdocalyptus dawsoni*) the tissues are very flimsy, consisting of thin, perforated sheets and filamentous strands draped around a scaffold of spicules, like a three-dimensional cobweb³. Few of the strands exceed 20 μm in thickness and the surface layers are about 1 μm thick.

To record from the sponge we dissociated tissue and allowed it to aggregate, using

Concanavalin A to promote adhesion^{4,5}. When the aggregates had rounded off, we placed them on the surface of slabs of body wall taken from the original sponge. After 12 h the grafts were attached and after 24 h cytoplasm could be seen streaming into the sponge (Fig. 1a). The tissue formed a substantial lump to which suction electrodes could be attached (Fig. 1b).

After 24 h, we put the slabs into a perfusion chamber in a shielded recording cage, inserted stimulating electrodes and positioned a thermistor-bead flow meter over the preparation to record changes in water flow across the body wall. We then attached

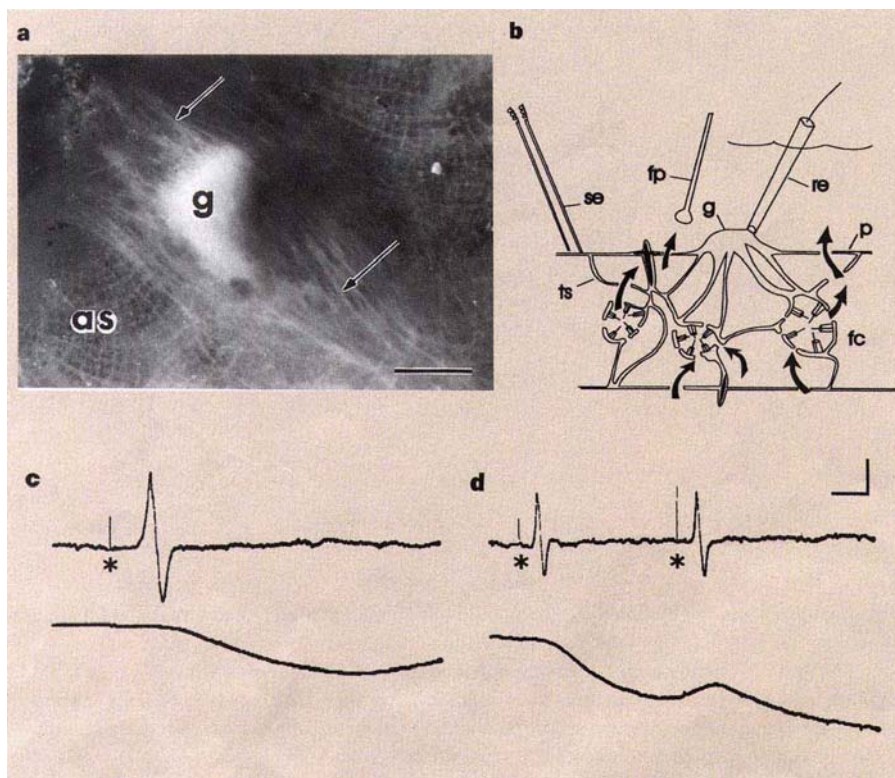


Figure 1 a, An attached graft (g) 24 h after being placed in contact with the atrial surface (as) of a piece of sponge. Arrows show streams of cytoplasm passing from the graft into the sponge. Scale bar, 1 mm. b, Diagram of a section through the body wall of *Rhabdocalyptus* showing a graft (g) fused to the pinacoderm (surface layer, p) of a piece of sponge. re, suction recording electrode; ts, trabecular syncytium; fc, flagellated chambers; arrows, water flow across the body wall; fp, probe to detect changes in flow velocity; se, chlorided silver stimulating electrodes. c, d, Recordings from such a preparation. Upper traces, electrical records from the suction electrode; lower traces, changes in water flow velocity past the flow-meter probe. Shocks (shown by asterisks) were delivered at a point 1.1 cm from the recording point. In c a single shock evoked a propagated action potential followed immediately by a reduction in water flow velocity. In d, a second shock delivered 40 s after the first evoked a second impulse and a more complete arrest of the feeding current. Time scale: 5 s in c, 10 s in d. Voltage scale (upper trace), 0.1 mV.

a polyethylene suction electrode (internal diameter 80 μm) to the graft, close to the flow probe. We stimulated the preparation with single electrical shocks, delivered more than 1.0 cm from the recording point, and displayed changes in water flow and associated electrical events, recorded through capacity-coupled amplifiers (0.1 Hz time constant), on an oscilloscope.

We consistently recorded an electrical impulse preceding each arrest of the feeding current (Fig. 1c). Amplitudes lay in the range 50–200 μV depending on the suction applied to the recording electrode. The signal was lost if the suction was insufficient to produce a good seal. Conduction velocities lay within the range of values previously reported for the spread of arrests (0.17–0.3 cm s^{-1} at 11 $^{\circ}\text{C}$)¹. After an impulse, the preparation was refractory for 30–40 s. A second shock 40 s after the first evoked an impulse identical to the first, with summing of the effector responses (Fig. 1d). Impulses were initially positive-going, showing a simple biphasic form and a duration of ~ 5.0 s.

We suggest that the tissue responsible for the conduction of the electrical activity is the trabecular syncytium, which makes up 75% of the organic matter in the sponge. This system penetrates all parts of the sponge, including the flagellated chambers where the effector response takes place. The syncytial character of this tissue, first reported using light microscopy, has now been confirmed by electron microscopy³ and *in vitro* studies^{4,5}. Aggregates are also syncytial⁶ and fuse with the trabecular syncytium when placed in contact with it. The absence of membrane barriers within the syncytium presumably allows action currents to spread in any direction from sites of depolarization. Thus, there is a clear path for the propagation of impulses through into the graft.

As yet, we do not know the ionic basis of the action potential, and although channels have never been identified in sponges⁷ we assume that voltage-activated channels are involved, as in other excitable tissues. Our findings place *Rhabdocalyptus* firmly on the list of animals capable of non-

nervous conduction⁸ and strengthen the view that excitability arose in the metazoa before the evolution of nerves.

Sally P. Leys, George O. Mackie

Biology Department, University of Victoria,
Victoria, British Columbia V8W 3N5, Canada
e-mail: mackie@uvic.ca

1. Lawn, I. D., Mackie, G. O. & Silver, G. *Science* **211**, 1169–1171 (1981).
2. Mackie, G. O., Lawn, I. D. & Pavans de Ceccatty, M. *Phil. Trans. R. Soc. Lond. B* **301**, 401–418 (1983).
3. Mackie, G. O. & Singla, C. L. *Phil. Trans. R. Soc. Lond. B* **301**, 365–400 (1983).
4. Leys, S. P. & Mackie, G. O. in *Sponges in Time and Space* (eds van Soest, R. W. M., van Kempen, T. M. G. & Brackman, J.-C.) 417–423 (Balkema, Rotterdam, 1994).
5. Leys, S. P. *Biol. Bull.* **188**, 241–254 (1995).
6. Pavans de Ceccatty, M. *Dev. Comp. Immunol.* **6**, 15–22 (1982).
7. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA, 1992).
8. Anderson, P. A. V. *Progr. Neurobiol.* **15**, 161–203 (1980).

Inferring seal populations from lake sediments

An explosion in the population of Antarctic fur seals (*Arctocephalus gazella*) has caused widespread changes to many coastal terrestrial and freshwater ecosystems in the northern maritime Antarctic islands and on the west coast of the Antarctic Peninsula. We have used seal hairs found in lake sediment cores from one maritime Antarctic island as a historical record of seal populations. This has enabled us to examine possible causes of the increasing numbers of visiting Antarctic fur seals, and has provided a historical framework from which to evaluate conservation plans to minimize the adverse effects of seals at sites of particular ecological significance.

Although hunted to near extinction during the nineteenth and early twentieth century, the numbers of Antarctic fur seals have recently been increasing^{1,2}. On Signy Island in the South Orkney Islands, the number of seals visiting each summer from the main breeding beaches on South Georgia has increased from less than 100 before 1976 to almost 20,500 in 1994 (ref. 2). This large increase has caused extensive destruction of vegetation, soil erosion and the eutrophication of freshwater lakes on coastlines where the seals haul out^{2–6}.

The Protocol on Environmental Protection was adopted by the Antarctic Treaty Nations in 1991 and it is expected to enter into full international force in the near future. Given the commitment of the Treaty Nations to limit adverse impacts on Antarctica, the current Antarctic fur seal population explosion raises three questions. First, is the increase a result of human or natural influences? Second, does the increase exceed the range of normal population variability during the past several thousand years? Third, how might any control measures that are deemed necessary be best implemented

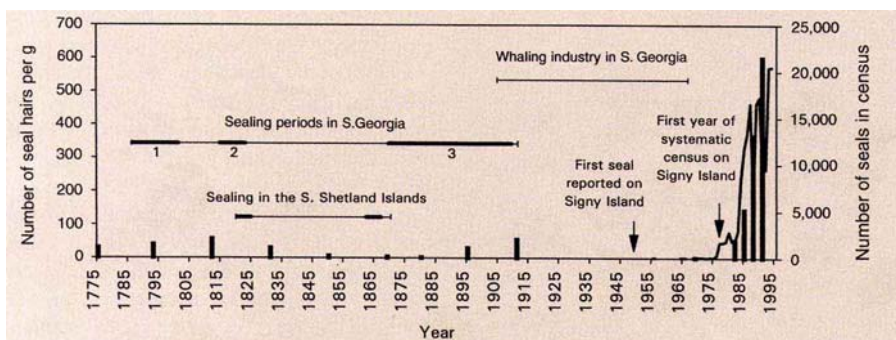


Figure 1 Histogram showing the number of Antarctic fur seal hairs in the upper 10 cm (AD 1775 to the present) of a short sediment core from Sombre Lake, Signy Island and line graph giving seal census data from 1977. Horizontal bars indicate periods when the sealing industry was active in the nearby South Shetland Islands and South Georgia, and when the whaling industry was active at South Georgia. Bold sections correspond to periods of peak harvest. Cores were dated with reference to an established radionuclide (^{210}Pb) and radioactive isotope (^{137}Cs) chronology¹⁰.

within the Antarctic Treaty system?

Answering these questions is not easy because of the lack of historical records. However, seal hairs found in lake sediment cores from Signy Island have allowed us to study the timing of the Antarctic fur seal population explosion in relation to possible causal factors. Geophysical and biological palaeoclimate indicators in sediment cores (V. Jones, personal communication) show that the population explosion is unlikely to have been influenced by changing natural ecological conditions⁷. However, the evidence does suggest a link between seal populations and the activities of the sealing and whaling industries.

A short sediment core showed that seals visited the island before commercial sealing began, but declined in numbers during sealing periods in the nearby South Shetland Islands between 1820 and the 1870s (Fig. 1). As sealers withdrew from the South Shetland Islands there seemed to be a brief increase in the abundance of visiting seals, but regional sealing based in South Georgia continued to reduce the population to near extinction. There is no evidence that fur seals visited Signy Island between 1950 and the late 1970s, and it was not until 1977 that the summer influx of seals began to increase rapidly. One possible cause of this increase might be the greater than 90% reduction in the number of baleen whales in the Southern Ocean by the whaling industry since 1922 (ref. 8). This seems to have resulted in an abundance of the Antarctic krill (on which whales feed) which has subsequently been available to seals^{3,9}.

We find from long sediment cores that the number of seal hairs presently being deposited in the sediments is 78–94% greater than at any time during the past 6,570±60 radiocarbon years, which implies that the present number of visiting seals exceeds the range of natural variability⁷. The number of seal hairs in sediment cores since 1977 corresponds well with census data (Fig. 1).

At present there is no evidence of any

particular species or community type being endangered by the impact of the seals, but in the northern maritime Antarctic islands and on the west coast of the Antarctic Peninsula, changes to lowland terrestrial and freshwater ecosystems in the past two decades are significant and are apparently spreading¹⁴. In the light of this, Antarctic Treaty Nations need to consider how ecologically important sites might be protected. The most practical option is physically to prevent fur seals from gaining access to areas of particular ecological or scientific importance, using wire mesh and electric fences. However, such proactive conservation management in Antarctica is controversial and may be rendered unnecessary when regulations set to manage stocks of whales (agreed by the International Whaling Commission), and stocks of fish and krill (regulated by the Convention on the Conservation of Antarctic Marine Living Resources) begin to exert a regulatory influence on seal populations in the future.

Dominic A. Hodgson, Nadine M. Johnston

British Antarctic Survey, High Cross,
Madingley Road, Cambridge CB3 0ET, UK
e-mail: daho@pcmail.nerc-bas.ac.uk

- Lewis Smith, R. I. in *Foundations for Ecological Research West of the Antarctic Peninsula*. Antarctic Res. Ser. 70 (eds Ross, R., Hofmann, E. & Quetin, L.) 15–59 (Am. Geophys. Union, Washington DC, 1996).
- Lewis Smith, R. I. in *Antarctic Communities: Species, Structure and Survival* (eds Bataglia, B., Valencio, J. & Walton, D.) (Cambridge Univ. Press, in the press).
- Lewis Smith, R. I. *Biol. Conserv.* **45**, 55–72 (1988).
- Lewis Smith, R. I. in *Antarctic Ecosystems, Ecological Change and Conservation* (eds Kerry, K. R. & Hempel, G.) 32–50 (Springer, Berlin, 1990).
- Croxall, J. P. et al. in *Antarctic Ecosystems, Ecological Change and Conservation* (eds Kerry, K. R. & Hempel, G.) 237–240 (Springer, Berlin, 1990).
- Ellis-Evans, J. C. in *Antarctic Ecosystems, Ecological Change and Conservation* (eds Kerry, K. R. & Hempel, G.) 77–82 (Springer, Berlin, 1990).
- Hodgson, D. A. et al. *Biol. Conserv.* (submitted).
- Headland, R. K. *The Island of South Georgia* (Cambridge Univ. Press, 1984).
- Jouventin, P. & Weimerskirch, H. in *Antarctic Ecosystems, Ecological Change and Conservation* (eds Kerry, K. R. & Hempel, G.) 208–213 (Springer, Berlin, 1990).
- Appleby, et al. *J. Paleolimnol.* **13**, 179–191 (1995).

Salty brines on the Mediterranean sea floor

In this communication, we report on brines from the Discovery Basin in the eastern Mediterranean which have the highest salinity ever found in the marine environment. They were formed by the dissolution of bischofite ($\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$) and filled the basin during the past 2,000 years. The brines give the first clear evidence for large-scale bischofite formation during the Messinian salinity crisis¹ (late Miocene epoch, 5–6 million years ago) and demonstrate that the eastern Mediterranean became evaporated to near dryness during this time.

The Discovery Basin was identified, named and sampled during a cruise of RRS *Discovery* between December 1993 and January 1994 (ref. 2). It is located on the western part of the Mediterranean Ridge (35° 20' N, 21° 40' E) at a depth of 3,580 m below sea level. The ridge is an accretionary complex situated at the convergent plate margin between the African plate and the Aegean arc³. The brine lake that fills the basin has a surface area of about 7.5 km² and a volume of nearly 2×10^8 m³.

Two sediment cores, PC15 and PC17, were taken at the southern and northern end of the Discovery Basin, respectively. Analysis of pore waters and water samples from the brine pool show that the brine is essentially a concentrated solution of MgCl_2 (Table 1).

The pore-water values in both cores are almost identical (Fig. 1), indicating that the basin is filled with a homogeneous brine throughout. Concentrations of Mg^{2+} and Cl^- decrease with sediment depth. The shape of the concentration profiles indicates that dissolved MgCl_2 is penetrating the sediments from above by both advection and molecular diffusion. We calculated the time taken to generate the measured profiles using a simple one-dimensional transport model⁴. Our calculations indicate that the surface sediments have been in contact with MgCl_2 brine for only 2,000 years. Within this short period the Discovery Basin must have been filled at an average rate of almost 10^5 m³ yr⁻¹.

During the evaporation of sea water, insoluble calcium minerals are precipitated first, followed by halite (NaCl), potash salts and finally bischofite, which is the most soluble of all marine evaporite salts⁵. Primary brines are formed by evaporation of water and secondary brines by the dissolution of evaporites. Equilibrium calculations with the PHRQPITZ model⁶ indicate that a primary brine similar to the Discovery brine may be produced when sea water is evaporated close to the point of bischofite precipitation (Table 1). A similar solution may also be formed when sea water is

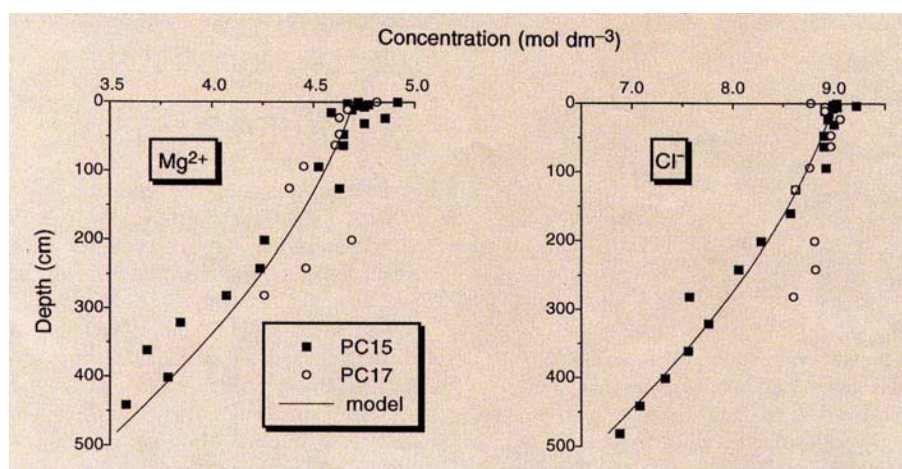


Figure 1 Pore-water profiles measured in cores PC15 and PC17. The model curves for Mg^{2+} and Cl^- are produced using the following parameter values: diffusion coefficient of the dissolved compound in the sediment, D_s , $2.4 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$; advection rate, V , 0.3 cm yr^{-1} ; time, t , 2,000 yr. D_s is based on the mutual diffusion coefficient for MgCl_2 valid for a 4.5 mol dm^{-3} solution of MgCl_2 (refs 12,13). It was corrected for temperature and tortuosity effects using the Stokes–Einstein equation and an empirical relation between porosity and tortuosity¹⁴.

equilibrated with bischofite and kainite ($\text{K}_2\text{Mg}_2\text{Cl}_4(\text{SO}_4)_4 \cdot 11\text{H}_2\text{O}$). During this reaction halite and gypsum ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$) precipitate from the brine because of the high dissolved Cl^- and SO_4^{2-} concentrations produced by evaporite dissolution. Na^+ and Ca^{2+} concentrations in the brine are thus lower than in normal sea water. Late-stage evaporites other than bischofite would produce secondary brines with higher concentrations of Na^+ , Ca^{2+} or K^+ than observed.

The major ion composition of the Discovery brine is consistent with either a primary or secondary origin, but trace element concentrations can be used to differentiate between primary and secondary brines because Br, B and Li do not form evaporite

minerals during seawater evaporation⁵. The values listed in Table 1 are calculated for a primary brine evaporated to the point of bischofite precipitation, taking into account the solid solution of Br^- in precipitated chloride minerals⁷. The low Li and B values of the Discovery brine indicate that the brine is not formed by seawater evaporation. Li^+ is conserved during evaporation and brine storage as it is neither co-precipitated with evaporites nor adsorbed on clays and other solids in the presence of high Mg^{2+} concentrations⁵.

The Br concentration may be produced by the dissolution of bischofite with a Br content of 8 mg g^{-1} . This Br content is characteristic for bischofite precipitated from an extremely concentrated brine with a Br con-

tent of 12 mg g^{-1} (ref. 8). Such large Br concentrations are only obtained when sea water is evaporated to less than 1% of the initial water volume. The formation of late-stage evaporites and brines is often associated with hydrothermal fluids at spreading ridges⁹, which are enriched in ^3He . The low radiogenic $^3\text{He}/^4\text{He}$ ratio measured in the brine clearly excludes such a hydrothermal origin of the Discovery brine. The high ^4He concentration indicates that the brine was stored in uranium- and thorium-bearing sediment layers before it entered the Discovery Basin.

In conclusion, the chemistry of the Discovery brines shows that extreme evaporite conditions in the eastern Mediterranean, probably during the late Messinian, resulted in the precipitation of bischofite, which was subsequently redissolved and has migrated to form a brine pool on the sea floor.

Klaus Wallmann, Erwin Suess

GEOMAR Forschungszentrum für Marine Geowissenschaften der Universität Kiel, Wischhofstrasse 1–3, D-24148 Kiel, Germany
e-mail: kwallmann@geomar.de

Graham H. Westbrook

School of Earth Sciences, University of Birmingham, Birmingham B15 2TT, UK

Gisela Winckler

Institut für Umweltphysik der Universität Heidelberg, Im Neuenheimer Feld 366, D-69120 Heidelberg, Germany

Maria B. Cita

Università Degli Studi Milano, Dipartimento Di Scienze Della Terra, Via Maniagalli 34, Milano, Italy
and the MEDRIF consortium

1. Hsu, K. J., Cita, M. B. & Ryan, W. B. F. *Init. Rep. DSDP 13*, 1203–1231 (1973).
2. MEDRIF consortium *EOS* 76, 313 (1995).
3. Lallemand, S. *et al. Tectonophysics* 234, 3–52 (1994).
4. Ogata, A. & Banks, R. B. *USGS Prof. Pap.* 411-A (1961).
5. Sonnenfeld, P. *Brines and Evaporites* (Academic, Orlando, 1984).
6. Plummer, L. N., Parkhurst, D. L., Fleming, G. W. & Dunkle, S. A. *USGS Water Resources Investigations Rep.* 88–4153 (1988).
7. Fontes, J. C. & Matray, J. M. *Chem. Geol.* 109, 149–175 (1993).
8. Braitsch, O. *Salt Deposits. Their Origin and Composition* (Springer, Berlin, 1971).
9. Hardie, L. A. *Am. J. Sci.* 290, 43–106 (1990).
10. Bayer, R. *et al. Sitzungsberichte der Heidelberger Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse* 5 241–279 (1989).
11. Pytkowicz, R. M. *Equilibria, Nonequilibria and Natural Waters* (Wiley, New York, 1983).
12. Miller, D. G., Rard, J. A., Eppstein, L. B. & Albright, L. G. *J. Phys. Chem.* 88, 5739–5748 (1984).
13. Felmy, A. R. & Weare, J. H. *Geochim. Cosmochim. Acta* 55, 112–131 (1991).
14. Lerman, A. *Geochemical Processes. Water and Sediment Environments* (Wiley, New York, 1979).

erratum

In Fig. 1 of "Biochemical typing of scrapie strains" (*Nature* 386, 564; 1997) by Robert A. Somerville *et al.*, which plots the glycoform profiles of several scrapie strains, the y-axis was labelled incorrectly. It should have read 'PrP^{Sc} in low M_r glycoform' rather than 'PrP^{Sc} in high M_r glycoform'. Also note that like-coloured shading does not imply any grouping of strains.

Table 1 Composition of the Discovery brine, of sea water, and of primary and secondary brines

Species concentration (mol per kg H ₂ O)	Discovery brine	Seawater composition (salinity 35‰) ^a	Primary brine A, H, Ki, Car [*]	Secondary brine Bi, Ka, H, G [†]
Na	0.084	0.4855	0.166	0.099
K	0.090	0.01058	0.071	0.112
Ca	0.001	0.01066	0.001	0.003
Mg	5.15	0.055	5.41	5.48
Cl	10.15	0.5658	10.1	10.9
SO ₄	0.110	0.02926	0.173	0.122
pH	4.5	~8.0		
ΣCO ₂	0.004	~0.0027		
Br	0.110	8.7×10^{-4}	0.110	
B	2.4×10^{-3}	4.5×10^{-4}	4.83×10^{-2}	
Li	7×10^{-4}	2.9×10^{-5}	3.35×10^{-3}	
He	6×10^{-8}	4×10^{-8}		
³ He/ ⁴ He	1×10^{-7}	1.4×10^{-6}		

Species concentrations were determined using flame emission spectrometry (Na, K), ICP spectrometry (Ca, Mg, B), argentometric titration (Cl), ion-chromatography (SO₄, Br), graphite-furnace AAS (Li), gas-chromatography (ΣCO₂), pH electrodes and a special sector field mass spectrometer¹⁰ (He, ³He). He concentrations are given in cm³ per g brine at 1 atm. The molal concentrations (mol per kg H₂O) in the Discovery brine are calculated as the molar concentration $\times (d-s)^{-1}$, where d is the brine density (1.33 kg dm^{-3}) and s the dissolved solids concentration in the brine (0.44 kg dm^{-3}).

^{*}Brine produced by the evaporation of sea water and precipitation of anhydrite (A, CaSO_4), halite (H), kieserite (Ki, $\text{MgSO}_4 \cdot \text{H}_2\text{O}$), and carnallite (Car, $\text{KMgCl}_3 \cdot 6\text{H}_2\text{O}$). In the PHRQPITZ⁷ calculation, a temperature of 30 °C was assumed and the evaporation was continued until only 5 g of the initial 1 kg H₂O remained in solution. The trace element concentrations are taken from ref. 7.

[†]Brine produced by equilibrating sea water (salinity, $S=35\text{‰}$) with the evaporite minerals bischofite (Bi), kainite (Ka), halite (H), and gypsum (G) at 14 °C. During the reaction, 5.322 mol Bi and 0.025 mol Ka per kg H₂O are dissolved and 0.386 mol H and 0.008 mol G per kg H₂O are precipitated, until equilibrium is established.

When the sky fell in on the dinosaurs

T. rex and the Crater of Doom

by Walter Alvarez

Princeton University Press: 1997. Pp. 185. \$24.95, £18.95

Clark R. Chapman

While the theories of plate tectonics were being developed in the 1960s, the Apollo missions revealed the ubiquitous planetary cratering that would eventually lead to another revolution in geology — the recognition that impacts by asteroids and comets have helped to shape the Earth's surface and influenced the evolution of life. Walter Alvarez contends in this slim book that the excitement of plate tectonics distracted most geologists from the lessons coming from space. Not until 1980, when Alvarez, his father and some colleagues published their paper proposing an extraterrestrial cause for the Cretaceous/Tertiary (K/T) mass extinctions, were geologists forced to confront cosmic catastrophism.

It has been a painful conversion for most geoscientists, because impacts were considered to be antithetical to the uniformitarian dogma that had dominated geology for two centuries, including the view that climate changes or a fall in sea level led to the mass extinctions. The tale of the heated debates in the 1980s about the Alvarez hypothesis is not fully told in this book, despite the promise on the dust cover that it recounts "intense public debate [and] friendships made or lost...". Although he writes of friendships made during field trips in dis-



...creating a giant sea wave that sweeps the dinosaurs to their doom.



Moment of impact: a comet or asteroid at least 10 km wide hits the Earth...

tant lands, Alvarez is too civil to delve deeply into the rivalries, media mistakes, public antagonisms or friendships lost.

The book has a simpler, yet fascinating story to tell — that of how Alvarez's geological career led him to study the K/T boundary layer at Gubbio, Italy, and how the immense Chicxulub crater was located in the Yucatán Peninsula, Mexico, seeming to prove his father's idea that global dust from an impact led to the holocaust 65 million years ago. The tale is personal, well written and delightful. Alvarez also reflects on the history and philosophy of geology, making this book well suited to geology students and literate non-geologists interested in how geologists think.

Alvarez portrays a modern geology that builds on the interdisciplinary resources of physics, chemistry and astronomy to fashion a synthesis about how our planet has evolved.

Discovery of the famous iridium anomaly in the K/T boundary layer appears part of the way through the book, beginning the mystery story about the hunt for the "crater of doom". Although he gives credit where credit is due, this is primarily the story of Alvarez's own quest for the answer, involving both miscues and successes, finally focusing on the 65-million-year-old Mexican crater.

We read of tsunami deposits in Texas and Haiti, geophysical identification of the buried crater in the Yucatán, discovery of microtektite spherules in the boundary layer, and the discovery and dating of melt

rocks. Strangely, Alvarez's personal turning point came later, in 1992: he tells of a moment of epiphany in the central plaza of a small Mexican town.

Sticking close to his own role and expertise, Alvarez omits other elements of the K/T story. His father's idea that the dinosaurs, at the top of the food chain, died from starvation because of the collapse of photosynthesis during year-long darkness is introduced but never analysed. He does not discuss (except in a prologue) the many other environmental consequences of impacts that have been researched during the past 15 years. Much of the antagonism of palaeontologists to the Alvarez idea that the dinosaurs died out suddenly is caused by their inability to reconcile ostensible killing mechanisms with the fossil record. Yet Alvarez does not choose to try to address their concerns here.

Towards the end of the book, Alvarez mentions without amplification a few unresolved issues, such as whether other mass extinctions in the geological record might have been caused by impacts. He closes with a brief reflection on the symbolism of the crash of comet Shoemaker-Levy 9 into Jupiter.

One can read *T. rex and the Crater of Doom* in a single sitting and I recommend it highly — if only as a jumping-off point to other perspectives on this dramatic scientific revolution. □

Clark R. Chapman is at the Southwest Research Institute, 1050 Walnut Street, Boulder, Colorado 80302, USA.

A physicist's portrait gallery

Atomic Histories

by Rudolf E. Peierls

American Institute of Physics Press/Oxford University Press: 1997. Pp. 378. \$32.95, £27.50

Nicholas Kurti

This is a collection of the writings of the physicist Sir Rudolf Peierls from 1940 until the last years before his death in 1995. The earliest of these is the Frisch–Peierls Memorandum, submitted to the UK government in March 1940, which established the feasibility of the atomic bomb and initiated the sequence of events that led to Hiroshima and beyond, and gave Peierls a lasting place in history — added to his pre-eminence in physics.

It is ironic that, although he and Otto Robert Frisch were excluded from secret war work because they were enemy and later “ex-enemy” aliens, they could speculate with impunity on the problem of the atomic bomb and so lay the foundation of one of the most secret projects.

The text of the memorandum is given in full: both the technical part, already included in Margaret Gowing’s *Britain and Atomic Energy*, and a very readable non-technical “executive summary”, previously published only in Ronald Clark’s *Tizard* and containing many wise and even prophetic remarks.

In the technical part there is a footnote prefaced by “Note added March 1994”. It refers to the fission cross-section of uranium-235, which is used for calculating the critical size, and says: “This was a surprising slip. The correct figure is about 3 times smaller and the critical mass is... about 30 times too small” (600 g instead of 20 kg). But this error was soon discovered and for instance Franz Simon’s *Report on the Uranium Isotope Separation Plant* (December 1940) envisaged a production rate of 1 kg of uranium-235 per day and the Maud Report (July 1941) assumed much smaller cross-sections.

It is therefore difficult to agree with the occasionally expressed view that this serious underestimate of the critical mass was largely responsible for the allies’ decision to go ahead with developing the bomb, while the Germans’ gross overestimate of the size explains their lack of effort.

Peierls had a thorough knowledge of the subject and understood the dangers involved from the very first. This is clear from articles such as “The Defence Against the Atomic Bomb” (1946) and “Atomic Energy: Threat and Promise” (1947), and later writings such as “Britain in the Atomic Age” and “Reminiscences of a British Participant”. He was a founding member and vice-president of the Atomic Scientists Association which, in the United Kingdom, optimistically put forward



Peierls: understood the threat of the bomb from the very first.

the idea of international control of atomic energy but was dissolved in 1959. Later he attended the Pugwash Conferences which concentrated on arms control and disarmament. These articles remind today’s readers that the questions of atomic energy, and of the bomb in particular, were always regarded with deep seriousness by the scientists concerned.

About half the book is taken up by “A Physicist’s Portrait Gallery” — accounts of the life and work of Peierls’ many friends and colleagues. He was born in Germany but he worked in many European countries, the United States and Russia (his wife was Russian). Not for nothing did he call his autobiography *Bird of Passage*. He was acquainted with the cream of twentieth-century physicists.

It is clear that Peierls regarded Werner Heisenberg as one of the greatest physicists of this century. Four articles have Heisenberg as their main subject, including a 20-page article for the *Biographical Memoirs of the Royal Society*. All these articles discuss Heisenberg’s role in Germany’s Atomic Energy Project and Peierls is meticulously fair to him when assessing the reasons for the half-hearted and somewhat ineffectual nature of the German work. But Peierls also criticizes the view that the German scientists knew how to make an atomic bomb but that moral scruples made them all but sabotage the project.

After being told that the Americans had dropped an atomic bomb on Hiroshima, Heisenberg said: “If it has been done with uranium-235 then we should be able to work it out properly. It just depends on whether it is done with 50 or 500 or 5,000 kilogrammes and we don’t know the order of magnitude. We can assume that they have some method of separating isotopes of which we have no idea.” (Quoted from the *Farm Hall Transcripts*.) He said this five years after the Frisch–Peierls recipe for the atomic bomb and Simon’s blueprint of a large-scale plant for isotope separation.

Other portraits in this gallery include those

of Niels Bohr, James Chadwick, Paul Dirac, Lev Landau, J. Robert Oppenheimer and Ernest Rutherford. Some of these are lengthy biographical memoirs written for the Royal Society, others are shorter appreciations, often reviews of the subjects’ biographies. The memoirs on Frisch and Wolfgang Pauli stand out by showing Peierls’ skill in blending the description of the subject’s scientific work with the portrayal of his personality. The scientific explanations are models of clarity and simplicity while the character sketches reflect Peierls’ affection for his subjects. This does not make him blind to their foibles, which are recognized and often made fun of, but the irony, although sharp, is never malicious.

There are many examples of Pauli’s merciless criticism of his students or colleagues. But, as Peierls says, “the remarkable thing is that, although the victims often felt hurt at the time, none of them bore a grudge for long”. Peierls might have mentioned the occasion when he himself was the subject of a savage attack by Pauli after a seminar talk. Pauli then turned to Peierls with a disarming smile and asked: “Would you not have said the same thing if I had given the talk?” Peierls replied: “Yes, but I would have said it more politely.”

Frisch was an experimental physicist of great ingenuity. It was he who suggested the term ‘fission’ when, with his aunt Lise Meitner, he correctly interpreted the Hahn–Strassmann experiment. Peierls’ portrait of Frisch shows him as a gentle, modest person, happy in his lab, happy with his physics, and happy with his music. He was a pianist of near professional standard who demonstrated the power of music over physical frailty. He was a bad sailor and yet I recall that, during a stormy Atlantic crossing in December 1943, when to be upright was agony, he gave a superb rendering of Chopin’s “Barcarole” — but had to be back in his bunk before the applause had died down.

This collection of some of the writings of one of this century’s most versatile and influential physicists makes wonderful reading and gives one a feel for nearly seven decades of physics and for its practitioners. The book is full of quotable passages, so let us finish with one from Peierls’ pen that shows his whimsical humour, his willingness to poke fun at himself and his utter lack of pomposity.

In 1955 Peierls published *The Laws of Nature*, a book intended for the non-specialist reader. It was dedicated “To the Treasurer of Somerville College, Oxford and the Bursar of Gonville and Caius College, Cambridge, but for whom this book would not have been written”. The treasurer of Somerville was puzzled by the dedication but the bursar of Caius saw the joke: Peierls wanted the royalties to help with the education of his daughter Gaby and his son Ronnie.

Nicholas Kurti is at the Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK.

Book of man

History of Physical Anthropology: An Encyclopedia

edited by Frank Spencer

Garland: 1997. 2 Volumes. Pp. 1,195. \$175

Clark Howell

This book is the first comprehensive encyclopaedic treatment of the discipline of physical anthropology — the natural history of humankind. It provides brief biographical sketches of some 302 figures, mostly historical; summaries of development, both disciplinary and institutional; and entries on essential areas of scientific enquiry and theory. There are 169 contributors, from 34 countries, and the masterful editorial hand of Frank Spencer is manifest throughout the consistently readable text. He contributes many biographical profiles and several topical entries.

There is an extensive index of names and a thoroughly cross-referenced subject index. Much of the history of the field is exemplified in the abundant biographical profiles. In the 1800s, it is represented by many physicians, anatomists, zoologists, naturalists, palaeontologists and geologists. At that time, some form of 'normal' science was practised, but the influence of Darwinism, even if acknowledged, was scarcely noticeable. Only in the past 50 years did the field become professionalized.

Some recent figures who were particularly important to the profession lack biographical treatment: W. C. Boyd, A. A. Dahlberg, A. E. Mourant, H. J. Muller, S. Wright, A. C.

Wilson and J. S. Huxley (whose grandfather does appear). Thirty-five entries summarize the emergence, development and modern condition of the field — largely since the 1950s — in countries of the Americas (10), Europe (16), Africa (1), Middle East (2), Asia (4) and Antipodes (2). Further insights into historical aspects are found in entries on scientific academies, professional societies and associations, laboratories, museums, universities, journals/serials and influential foundations and government bodies.

No single entry reflects the nature, scope and theory of physical anthropology. The discipline is not about *chain of being*, *catastrophism*, *creationism*, *eugenics*, *phrenology*, *polygenism*, *Rassenhygiene*, *Rassenkunde*, all of which have their own entries. The field does acknowledge and employ concepts, approaches and methodologies, including *biogeography*, (neo-) *Darwinism* (not *Lamarckism*), *evolutionary theory*, (*Mendelian*) *genetics* (and *genetic drift*), *punctuated equilibrium*, *paleoanthropological theory*, *race*, *systematics*, *taphonomy* and *uniformitarianism*. But there should have been entries as well on cladistics, ethnic groups, extinction, life history, phylogeny, populations, species and taxonomy.

Most concerns of the discipline encompass *adaptation*, *anthropometry* (as technique), *bipedalism* (within locomotion), *body composition*, *comparative neuroanatomy*, *demography* (and associated population biology and genetics), *dental anthropology*, *dermatoglyphics*, *growth* (and *aging*), *health and disease*, *molecular anthropology*, *paleoanthropology*, *primatology* (all aspects), *skeletal biology* (and *paleopathology*).

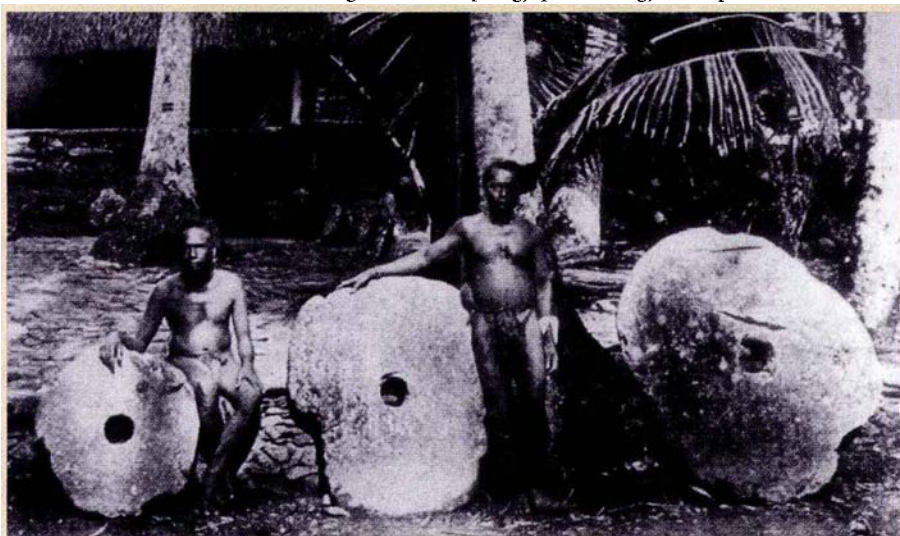
The encyclopaedia also deals with the interface and interactions between biology, behaviour and culture (as in the *hunting hypothesis of human origins*).

Primate studies have become a central focus in a field long dominated by concerns with mankind. Since Edward Tyson in the seventeenth century, various individuals have contributed knowledge of primate biology, diversity and structure. Field studies were minimal and pioneering until some 40 years ago. A half-dozen entries deal with the history and status of prosimian, simian and ape field studies in general and on several continents. Similarly, studies of the primate fossil record have expanded vastly in recent decades, and one of the longest and best entries is that on *paleoprimatology*.

Palaeoanthropology, reflecting human evolutionary studies from biological and cultural/behavioural perspectives, is among the most visible and familiar of the field's foci. Strangely, it lacks an inclusive entry (other than an historical treatment, *paleoanthropology theory*). It principally encompasses *Java*, *Asian* and *Australian* (*paleoanthropology*), and an expansive and topical piece on *modern human origins*. The relevant fossil record is treated idiosyncratically with entries on *Australopithecines*, *Homo habilis*, *Homo erectus* and *Neanderthals*. *Homo sapiens* is not to be found, although more than 20 index citations under *Homo* are to species nomina, among them *sapiens*, including the informal (and unacceptable) 'archaic' and 'early modern' designations. The lack of exposition of hominin systematics is a serious omission. Many, but by no means exhaustive, entries on important fossil-bearing formations/localities — from *Afar* to *Olduvai Gorge* and *Omo* to *Zhoukoudian* — provide additional information to flesh out the treatment of hominid (hominin) evolution and phylogeny. Lacking are entries on *Atapuerca*, *Dmanisi* and *Kanapoi*, among others. Fuller, more structured treatments are to be found in *The Cambridge Encyclopaedia of Human Evolution* edited by S. Jones, R. Martin and D. Pilbeam (Cambridge University Press, 1992; for a review see *Nature* 361, 314; 1993) and in *Encyclopaedia of Human Evolution and Prehistory* edited by I. Tattersall, E. Delson and J. van Couvering (Garland, 1988; see *Nature* 335, 598; 1988).

History of Physical Anthropology is an important addition to the reference literature. It will inform nonbiological practitioners about the study of humankind as well as those in the many increasingly fissioned specialities of the life sciences. □

Clark Howell is in the Laboratory for Human Evolutionary Studies, Museum of Vertebrate Zoology, University of California, Berkeley 94720, California, USA.



Hard currency

Nineteenth-century Yap islanders made big money. But there was no danger of these coins burning a hole in the pocket as they took the form of stones up to 4 metres across. The stones were cut from limestone quarries on islands some 400 miles away and carried by sea to Yap, in what is now Micronesia. From *Money*:

A History, edited by Jonathan Williams (British Museum Press, £25, St Martins Press \$29.95). The book was written by curators at the British Museum in London and was published to coincide with the opening of a new gallery at the museum exploring the history of money.

Travels with diseases

Who gave Pinta to the Santa Maria? Tropical Diseases in a Temperate Climate

by Robert S. Desowitz
Norton: 1997. Pp. 228. \$25

Len Goodwin

Bob Desowitz has spent many years in tropical countries as a researcher and teacher of parasitology. He is now a bearded patriarch with an emeritus chair in Honolulu, and from time to time he sets down his view of the past, present and future of some aspect of the science.

The parasites of animals developed and varied along with their hosts and, when the continents separated, those in the New World took different lines from those in the Old. So it is not surprising that the exchange of people and parasites that occurred during the age of discovery led to epidemics of disease that changed the course of history, the maps of the world and the fate of nations. Desowitz outlines some of the most exciting.

European sailors took home spirochaetes from America and started an epidemic of syphilis. The pinta of the book's title is a skin infection. Missionaries carried measles, smallpox and benign tertiary malaria to the New World, and the slaves from Africa added malignant malaria. The epidemics of unknown diseases, together with terror and despair, killed off most of the Amerindians.

Yellow fever had a profound effect. Desowitz describes the epidemics that enabled the United States to buy vast areas of land that the European powers could no longer support. He describes the work of Wal-

ter Reed's team that elucidated the transmission of the virus by mosquitoes. By targeting the species of mosquitoes that carry yellow fever and malaria, William Crawford Gorgas was able to complete the construction of the Panama Canal, which had been abandoned by the French.

There is an interesting description of the influence of the millionaire philanthropist John D. Rockefeller and the foundation he financed in clearing hookworm from the southern United States, helping to control malaria and dealing a final blow to yellow fever by the development of a safe vaccine.

Exotic disease is at present under control in temperate countries but new introductions are a constant danger, as shown by unexpected outbreaks, usually traceable to immigrants from endemic areas. Temperate climates can support the vectors of many tropical infections and, with global warming, are likely to become even more favourable. South American trypanosomiasis or Chagas' disease is carried by a bug; suitable species, together with reservoir hosts — ground squirrels and dogs — are widely distributed in the United States. It would not be too difficult for an epidemic to occur and there is no really effective treatment. Also, as the remaining tracts of wilderness are invaded, new infections such as Kyanasur Forest and Ebola viruses may be transmitted from animals to humans.

New drugs for exotic diseases are not a priority for the pharmaceutical industry, although the old cures are becoming less effective. A serious situation could develop before long. Desowitz makes the point that fundamental discoveries that have changed the face of medicine have often come from completely unrelated fields of science and he appeals for more support to be given to basic research with no immediate short-term targets.

The style is chatty, wry and racy — the book was probably derived from popular lectures or student seminars — and the narrative is broken by frequent philosophical digressions, personal recollections and anecdotes. There are a few terrible illustrations but a dearth of references and no index. It would make ideal, informative reading for a fairly long flight to a tropical country. □

Len Goodwin is at Shepperlands Farm, Park Lane, Finchampstead, Berkshire RG40 4QF, UK.

New journals

This year, *Nature's* annual new journals review supplement will appear in the issue of 11 September. Publishers and learned societies are invited to submit journals for review, as well as details of any eligible electronic journals, taking note of the following criteria:

- Journals that first appeared during or after June 1995 and issued at least four separate numbers by the end of May 1997 will be considered.
- Journals covering any aspect of science are eligible, although those dealing with clinical medicine and pure mathematics are excluded, as are publications of abstracts.
- Frequency of publication must be at least three times a year. The main language used must be English. Translation journals in English are, of course, eligible.
- Deadline for submission is 6 June.

When submitting journals for review, please send at least four different issues (the first, the most recent and any two others) of each title, together with full details of subscription rates. For further information please contact Peter Tallack, *Nature*, Macmillan Magazines, Porters South, Crinan Street, London N1 9XW, UK. Tel: +44 (0)171 843 4567. e-mail: p.tallack@nature.com.

New in paperback

Why Things Bite Back: Predicting the Problems of Progress

by Edward Tenner
Fourth Estate, £7.99

A look at the paradoxes and unintended consequences of technology. It "offers a much-needed healthy balance between contemporary technological utopian fantasies and neo-Luddite despair. This superb guide to our high-tech world deserves a wide readership", wrote Howard P. Segal in *Nature* 382, 504 (1996).

The Same and Not the Same

by Roald Hoffmann
Columbia University Press, \$22.50

A guide for the general reader to the art and science of chemistry. "Roald Hoffmann wants to tell readers not why chemistry is important but why he loves it", wrote Philip Ball in *Nature* 380, 34 (1996).

The Song of the Dodo: Island Biogeography in an Age of Extinction

by David Quammen
Simon & Schuster, \$17

David Quammen "summarizes in exciting, earthy prose what he learned about island biology and species extinction during eight years of adventurous travel, reading and conversation. Descriptions of places alternate with chapters on the history of ideas and accounts of conversations with biologists", wrote Lawrence B. Slobodkin in *Nature* 381, 205 (1996).

Altered States: Gene Therapy and the Retooling of Human Life

by Jeff Lyon and Peter Gorner
Norton, \$15.95, £11.95
Reviewed by David Weatherall in *Nature* 375, 545 (1995).

Grooming, Gossip and the Evolution of Human Language

by Robin Dunbar
Faber, £7.99
Reviewed by Derek Bickerton in *Nature* 380, 303 (1996).

The Artful Universe: The Cosmic Source of Human Creativity

by John D. Barrow
Penguin, £12.50; Little Brown \$17.95
Reviewed by Christopher Longuet-Higgins in *Nature* 379, 216 (1996).

The Neanderthal Enigma: Solving the Mystery of Modern Human Origins

by James Shreeve
Penguin, £8.99; Avon Books \$14
Reviewed by Jean-Jacques Hublin in *Nature* 381, 658 (1996).

Implications of crustal property variations for models of Tibetan plateau evolution

Thomas J. Owens* & George Zandt†‡

* Department of Geological Sciences, University of South Carolina, Columbia, South Carolina 29208, USA

† Department of Geosciences, University of Arizona, Tucson, Arizona 85721, USA

‡ IGPP, Lawrence Livermore National Laboratory

Shear-coupled teleseismic P waves sampling the interior of the Tibetan plateau provide evidence of systematic variations in crustal structure. The crust thins by up to 20 km from south to north with a concomitant increase in Poisson's ratio from normal values in the south to unusually high values in the north. This suggests that the crust of the northern plateau is partially melted due to high temperatures. These changes imply spatial and perhaps temporal variations in the way the elevation of the high plateau is created and maintained.

It is widely accepted that the high, uniform elevation of the Tibetan plateau¹ is, to a large extent, the result of crustal thickening in response to the collision of the Asian and Indian plates beginning 45–55 Myr ago. Although it has been known for at least a decade that the crustal thickness and crustal seismic properties are not uniform throughout the plateau^{2,3}, until recently our ability to resolve the details of these lateral variations was limited, so little effort has been made to seek a systematic explanation for these differences. Since 1991, a series of Sino-American^{4–6} and Sino-French^{7,8} seismological expeditions have gathered the first extensive data sets within the plateau itself, thus providing the opportunity to map crustal properties in greater detail.

We present here new observations of teleseismic shear-coupled P waves recorded by the 1991–92 Sino-American Tibetan plateau broadband experiment⁴ (Fig. 1) illustrating that significant differences in average crustal thickness and average Poisson's ratio exist within the high (>4.5 km elevation) Tibetan plateau. Our results indicate that the average P-wave velocity is low ($6.0\text{--}6.2\text{ km s}^{-1}$) and does not vary significantly in the region of the east-central plateau sampled by our observations. However, we also find that the crust is 10–20 km thinner in the northern plateau relative to the south, and that Poisson's ratio is much higher than normal in the northern plateau. Changes in crustal properties appear to be associated with changes in upper-mantle structure and properties. The northern plateau is underlain by upper-mantle material that is anomalous in many parameters^{3,9–11} that are indicative of high temperatures. Our observed variations in crustal properties strongly indicate that the crust is responding through thinning, partial melting, and perhaps lower-crustal flow. We conclude that the plateau has evolved by a combination of processes that accommodated India–Asia convergence, including significant underthrusting of India beneath the southern Tibetan plateau, and crustal shortening and flow in the north with lateral flow in the underlying mantle induced by collision of strong Indian lithosphere into weak Asian continental lithosphere.

Shear-coupled P waves in the Tibetan plateau

P waves generated from teleseismic and deep regional SV waves at major interfaces in the crust and at the free surface are frequently observable^{12–15}. Using the common notation¹⁶, the phases of interest in this study (Fig. 2) are Sp, an S-to-P conversion from the crust–mantle boundary (Moho) beneath the station; SsPmp, an S wave converted to a P wave at the free surface and then reflected up to the site at the Moho; and SsPmsPmp, a multiply reflected converted

phase from the Moho. These phases are most easily observed for deep earthquakes which have simple source functions and depth phases arriving late enough so as not to interfere with the shear-coupled P waves produced in the crust. The differential arrival times of Sp and SsPmp relative to the direct S wave (technically, Ss in our notation) can be used to estimate the bulk properties of the crust although, like most seismological observations, trade-offs exist between estimated parameters if amplitudes are not included (Fig. 3).

The shear-coupled P waves we discuss are from one deep, strikingly impulsive earthquake recorded during the 1991–92 deployment (92.087.20.03.07; ref. 4). We have observed similar phases for more than 15 events occurring in the western Pacific subduction zones during the deployment. However, the simplicity of this single event and the fact that it was clearly recorded at 10 sites make its analysis particularly convincing. It is immediately obvious from this event that there is significant lateral variation in crustal structure in the eastern Tibetan plateau (Fig. 2). Comparison of some adjacent stations show remarkable long-period similarity for the Sp and SsPmp phases, and often even for the SsPmsPmp phase. Comparison of the long-period waveforms indicates similarity of average structure of crust sampled at stations LHSA, SANG and GANZ. ERDO and BUDO have similar seismograms that are significantly different from the southern sites. Crustal-structure differences between these station groups is manifest in the change in arrival time with distance relative to the S-wave (the 'move-out') of the SsPmp and the SsPmsPmp phases (Fig. 2). It is clear from these move-outs that the crustal P-wave velocity, v_p , and/or the crustal thickness, h , are changing from north to south. However, the Sp phase does not move out relative to the S wave. This remarkable observation indicates that any changes in v_p and h are offset by changes in Poisson's ratio, σ .

We take a two-step approach to modelling these variations in the observed S waveforms. First, we measure $S - Sp$ and $SsPmp - S$ travel times at all stations. Using simple travel-time relationships, we can quickly examine a large number of combinations of h , v_p and σ to find the range of parameters that can match the observed times (Fig. 3). This initial analysis produces preferred models that separate the southern sites (clearly thicker crust and average Poisson's ratio at LHSA, XIGA, and, less clearly, at GANZ and SANG) from the northern sites (thinner crust with high Poisson's ratio at ERDO and BUDO) and produced intermediate parameters at AMDO and WND0 (Fig. 3). These trade-off curves demonstrate that a constant thickness crust with $h = 65\text{ km}$ and $\sigma = 0.27$ would

have to range in mean crustal velocity from 5.8 km s^{-1} (in the south) to 6.7 km s^{-1} (in the north) to match the move-out of the SsPmp phase. Such a large variation in crustal velocities is precluded by travel times of the regional-distance crustal phase, P_g (ref. 10). Within a more reasonable range of mean crustal velocities, the trade-off curves indicate that the range of thicknesses observed in previous studies could be accommodated by thick, slow, average-

Poisson's-ratio crust in the south (70 km, 6.0 km s^{-1} , 0.27) and thin, fast, high-Poisson's-ratio crust in the north (55 km, 6.4 km s^{-1} , 0.35). These estimates yield reasonable starting models for full reflectivity forward modelling¹⁷ in which we modified the details of the models in order to match the timing and amplitude of these phases in the broadband waveforms.

We present velocity models and waveform fits at SANG, WNDO

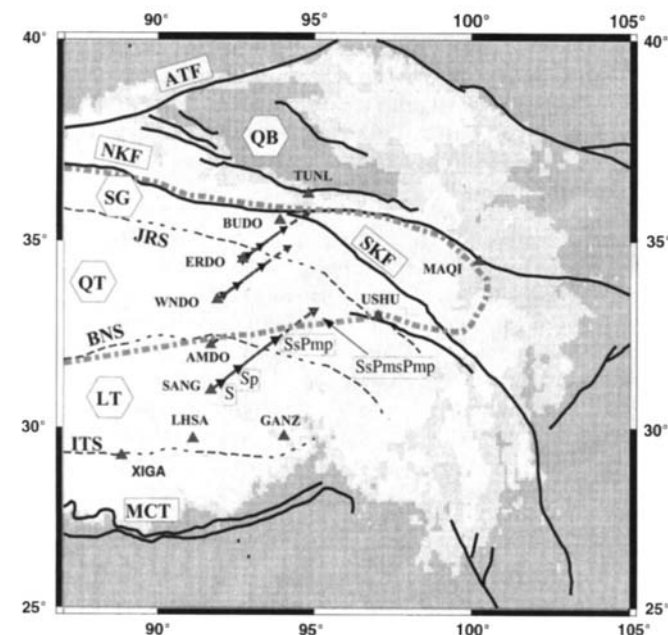


Figure 1 Map of the eastern Tibetan plateau. Stations of the 1991-92 Sino-American seismic experiment are named and shown by solid triangles. Extending from sites SANG, WNDO and ERDO to the northeast are lines connected by inverted triangles. The lines mark the paths of the shear-coupled P waves in the crust for event 92.082.20.33.07. As labelled at SANG, the inverted triangles mark the points at which the S wave and Sp, SsPmp and SsPmsPmp enter the crust. These reflections travel hundreds of kilometres in the crust and are thus sensitive to its average properties. The thick dashed grey line shows the extent of the region of inefficient S_n propagation in the upper mantle as identified by the 1991-92 Sino-American experiment¹⁰. Regional phase analysis from that study also showed that the upper-mantle P_n velocities were more than 3% lower north of 32°N compared to the southern plateau, and that crustal P_g travel times require the crust to be, on average, 15 km thinner north of 32°N (ref. 10). The 1993 Sino-French experiment produced evidence for an abrupt thinning of the crust north of the Jinsha River suture from offsets of regional phase arrivals and teleseismic phase analysis^{8,32}. Major accreted terranes are labelled: LT, Lhasa terrane; QT, Qiangtang terrane; SG, Songpan-Ganz terrane; QB, Qaidam basin. Major suture zones are indicated with thin dashed lines: ITS, Indus-Tsangbo suture; BNS, Banggong-Nujiang suture; JRS, Jinsha River suture. Major faults are shown in heavy black lines, each is labelled inside white boxes: ATF, Altyn Tagh fault; NKF, North Kunlun fault; SKF, South Kunlun fault; MCT, Main Central thrust. The darkest-grey areas are below 3,000 m elevation, the light-grey areas are between 3,000 and 4,000 m and the white areas are above 4,000 m elevation.

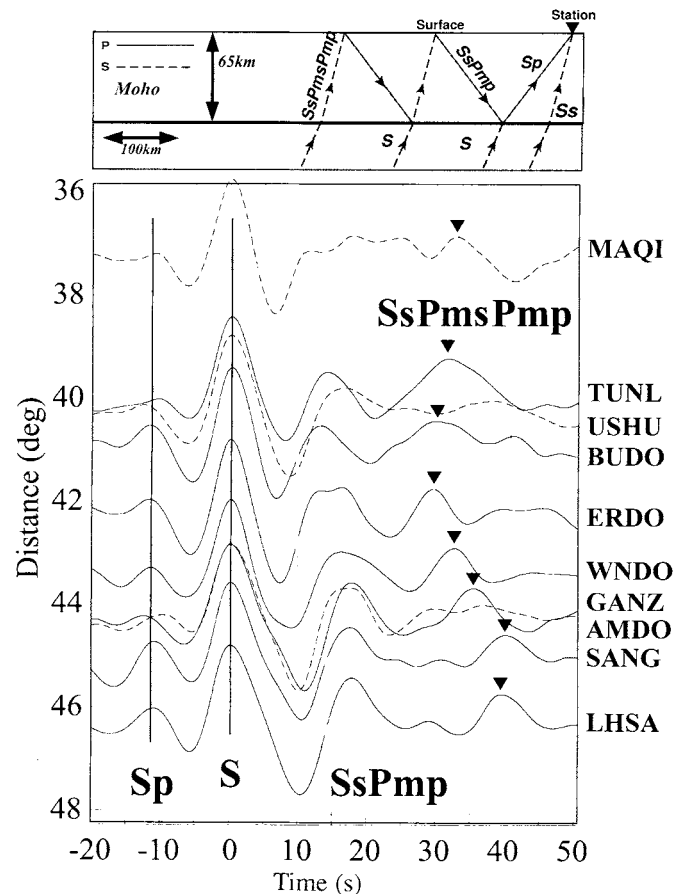


Figure 2 Top, ray paths of shear-coupled P waves in the crust. Arriving from below, teleseismic shear-coupled P waves are post-critical reflections at distances less than the critical distance ($\sim 50^\circ$ for Moho reflections). Bottom, 10-s low-pass filtered, vertical component displacement seismograms (solid lines) for event 92.087.20.33.07 recorded by stations across the Tibetan plateau. Each trace has been normalized to the amplitude of the direct S wave. All traces are aligned so that the direct S wave arrives at time zero. We measure differential travel times using the peaks of the S, Sp and SsPmsPmp phases and the zero-crossing of the phase-shifted SsPmp. (In theory, any phase with two crustal legs of its path as an S wave and three legs as a P wave could arrive at the same time as SsPmsPmp. The phase we observe in our seismograms has a P-wave particle motion and for simplicity we refer to it as SsPmsPmp.) The traces for MAQI, GANZ and USHU are dashed to indicate that they are offset from the main line of the experiment. Thick vertical lines at the S and Sp arrivals illustrate the lack of move-out with distance between these phases. Thick grey lines between 45° and 47° and between 41° and 43° mark the times of the zero-crossings of the SsPmp phases at the southern and northern sites, respectively, illustrating the large move-out between the northern and southern plateau sites for this phase. Inverted triangles mark the peak of the SsPmsPmp, which also shows move-out with distance that is significantly more than predicted for a homogeneous crust. Event 92.087.20.33.07 occurred on 27 March 1992, 20:28:14 GMT at 47.863°N , 147.127°E , 454 km depth, $M_b = 5.5$.

and ERDO (Fig. 4) to illustrate the crustal variations between the Lhasa, Qiangtang and Songpan-Ganz terranes. We model the full broadband waveforms because they best illustrate a significant difference between the Sp phases across our network. These broadband waveforms also have complexities that we have not tried to model in all their details. Where possible, we have used starting models with layering based on prior studies, but modified based on the trade-off curves in Fig. 3. We then concentrate our modelling on small variations to these models necessary to match the amplitude and timing of the Sp, SsPmp and SsPmsPmp. We found that the additional constraint of the SsPmsPmp phase timing and amplitude reduces the range of possible models allowed by the Sp-SsPmp trade-off curves.

Modelling of the seismogram at SANG, which samples the northernmost Lhasa terrane, results in a crust that is 74 km thick. The average v_p value for the entire crust is 6.0 km s^{-1} and Poisson's ratio is 0.27, near the global average¹⁸. In addition, there is clear evidence for a high-velocity layer of thickness approximately 14 km at the base of the crust. It is the top of this layer with P-wave velocity 7.2 km s^{-1} that generates the Sp_{HVL} arrival (Fig. 4a) that is prominent at all southern sites. This layer is visible in models derived from receiver function analysis, resembling most closely the lower-crustal model at XIGA¹⁹. This layer also includes a strong positive gradient which reduces the net velocity contrast at the Moho discontinuity and is required to match the Sp amplitudes.

The simplest crustal model is at WNDO which samples the Qiangtang terrane. Below the low-velocity surface layer, the crust is of uniform velocity to 54 km depth. The lower crust consists of a strong positive gradient from 6.3 to 7.0 km s^{-1} in the bottom 10 km of the crust. This gradient must be at least 10 km thick to reduce the amplitude of the synthetic Sp phase to match our observations. Our tests indicate that gradients more than 15 km thick decrease the amplitude of the Sp too much. Total crustal thickness is 65 km, the average v_p is 6.1 km s^{-1} , and the average Poisson's ratio is a rather

high value of 0.29. The crustal thickness and homogeneity of the crust is consistent with results of receiver function analysis^{19,20}.

Shear-coupled P waves arriving at ERDO only sample approximately 50–150 km north of WNDO (Fig. 1); however they do sample the Songpan-Ganz terrane and the derived crustal structure is quite different (Fig. 4b). At ERDO, the SsPmp phase is 1.4 s earlier, while the Sp phases arrive at the same time relative to the direct S wave. Previous studies have established that the crust at ERDO includes a significant low-velocity zone in the mid-crust²¹. Our waveforms are only weakly sensitive to the upper boundary of this layer, so we fixed the depth, overlying P-wave velocity and the S-wave velocity contrast of this interface to previously determined values²¹. In order to match our waveforms, we need very low average S-wave velocities in the crust. Our modelling indicates that this can be obtained by increasing Poisson's ratio in the crust from the top of this low-velocity layer to the Moho. The final model includes a gradational lower crust similar to that inferred at WNDO. The resulting velocity structure has an average v_p of 6.1 km s^{-1} . However, the estimated crustal thickness is only 55 km and the average Poisson's ratio is a very high 0.33. Below the top of the crustal low-velocity zone, Poisson's ratio averages 0.35. These values are due to the very low shear velocities that we interpret as strong evidence for crustal melting.

Implications for plateau evolution

Our seismic results for the Tibetan plateau crust are combined with the INDEPTH interpretation for the Tethyan Himalayas⁵ and constraints on mantle structure (based mainly on results from the 1991–92 Sino-American deployment^{9,10,22,23} and the 1993 Sino-French deployment⁸) in an interpretative north–south cross-section across the eastern side of the Indo-Asian collision zone (Fig. 5). This interpretation is based on the best estimates of lithospheric-scale structure of the Tibetan plateau available to date.

Our present study's determination of a systematic decrease in

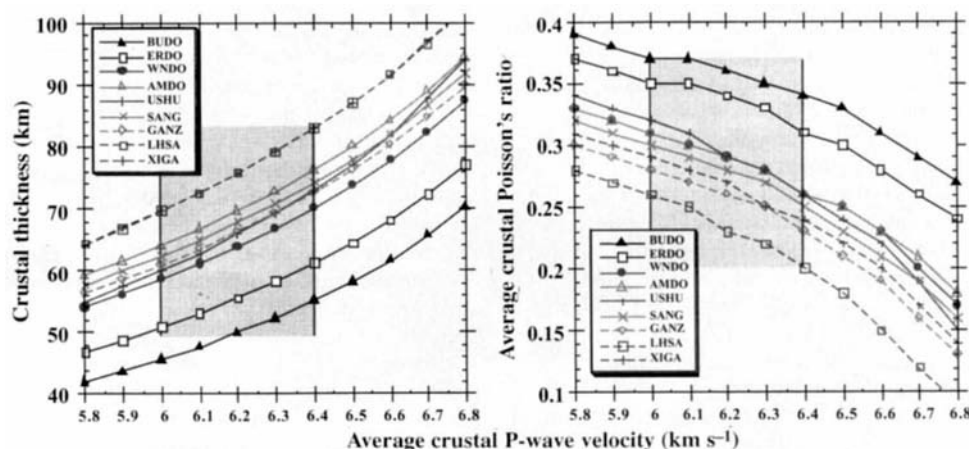


Figure 3 Trade-off curves for shear-coupled P waves in the crust of the Tibetan plateau. These plots were generated using simple, flat-Earth, single ray parameter travel-time equations for a one-layer model. We use the differential travel times measured at each station and the ray parameter appropriate for that station's distance. These times can be used to estimate the bulk properties of the crust. Ss and SsPmp both have one propagation leg through the crust as an S wave, thus their differential travel time depends only on the average P-wave velocity and crustal thickness. However, the differential travel time between Ss and Sp has an additional dependence on Poisson's ratio in the crust because the crustal leg of each phase is a different wave type. Here we parametrize the crust by its average thickness, h , average P-wave velocity, v_p and average Poisson's ratio, σ . Conversion of P-wave velocity and Poisson's ratio to S-wave velocity, v_s , can be made using the expression: $v_p/v_s = [(\sigma - 1)/(\sigma - 0.5)]^{1/2}$. Left panel, trade-

off between v_p and h determined from the SsPmp – S differential travel time. This panel shows clear differences between the southernmost and northernmost stations. These curves are satisfied for any value of σ . Right panel, trade-off between v_p and σ generated from the ratio of S – Sp time to the SsPmp – S time. This also clearly separates the acceptable models for the northern and southern sites. Note that each point on a stations $v_p - \sigma$ curve requires a different value of h to satisfy the travel-time equations. The vertical lines on each panel indicate the range of crustal velocities consistent with regional P_g analysis. The light-grey area between these lines highlights the range of possible crustal parameters implied by P_g velocity constraints¹⁰ and these trade-off curves. Curves for XIGA were generated from times measured for event 91.281.03.36.46; all others were generated using event 92.087.20.33.07.

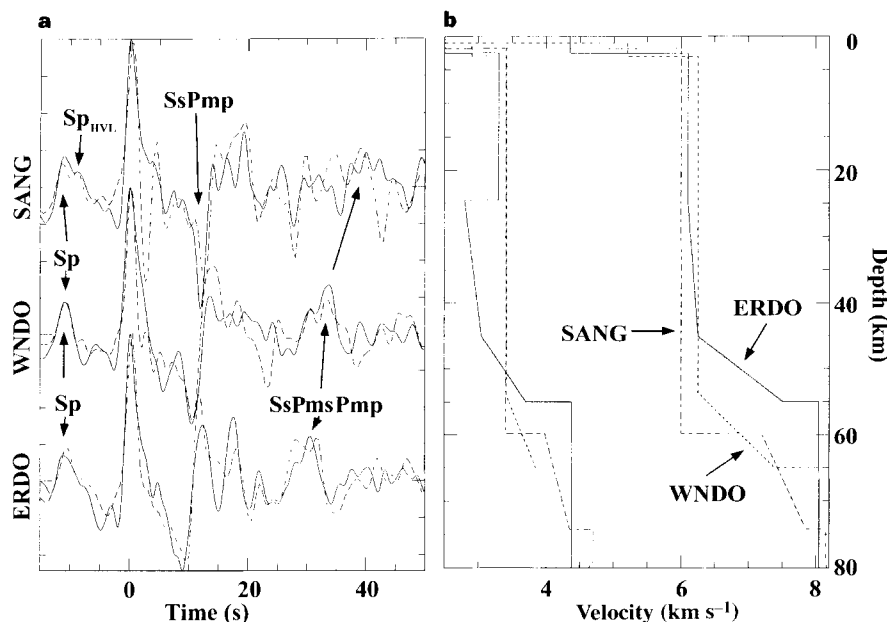


Figure 4 Synthetic seismogram modelling results for event 92.08720.33.07 at three sites, with derived crustal models. **a**, Broadband vertical component displacement seismograms (solid lines), with vertical component reflectivity synthetics (dashed lines). Our modelling concentrates on matching the timing and amplitude of the three labelled phases. Note the systematic decrease in $S - SsPmp$ time from south (SANG) to north (ERDO). **b**, Derived P- and S-wave velocity models for each site. All sites are characterized by nearly homogeneous P-wave velocity profiles. SANG includes an additional layer at the base of the crust whose sharp upper boundary generates the Sp_{HVL} arrival shown in **a**. WND and ERDO both have strong gradients at the base of the crust. The low

shear velocities in the crust of ERDO are responsible for the very high Poisson's ratio at that site. Each site has a low-velocity surface layer that both lowers the average crustal velocities and accounts for much of the 'ringing' in the broadband waveforms. Complex shallow structure at SANG probably results in the waveform mismatch ~ 2 s after the direct S wave for this structure. Our model testing suggests that we could match the data only by varying our preferred models by less than ± 5 km in h , less than $\pm 0.2\ km\ s^{-1}$ in v_R and less than ± 0.02 in σ by trading these parameters off against each other. However, larger variations prevented matching the $SsPmsPmp$ arrivals.

crustal thickness and the concomitant increase in Poisson's ratio have important implications for plateau evolution. These implications require us first to interpret the meaning of changes in Poisson's ratio. Values of this ratio between 0.25 and 0.30 can have a variety of compositional implications²⁴. In the initial stages of crustal melting, it is not clear how small percentages of melt will change Poisson's ratio. However, as Poisson's ratio is 0.5 in a fluid, it will necessarily increase as the volume of crustal melt increases. Unfortunately, we know of no studies relating the degree of melting to this inevitable increase in Poisson's ratio. Evidence suggests that unmelted crustal rocks with Poisson's ratio greater than 0.30 are rare²⁴. Thus, we believe that Poisson's ratios of 0.33 and higher in the northern plateau are strong evidence for extensive crustal melting.

We suggest (as have others) that the Indian continental lithosphere has underthrust Asia to roughly the Banggong suture²⁵, based partly on the observation of the high-velocity lower crust at stations LHSA, XIGA, GANZ and SANG. We interpret this layer as a signature of the Indian cratonic crust. The P-wave velocity of this layer ($7.2\text{--}7.5\ km\ s^{-1}$) is best interpreted as rocks of mafic or perhaps intermediate composition at the high pressure present at 60–74 km depth. In either case, associating this layer with the lower crust of underthrust Indian lithosphere also provides the simplest explanation for the high-velocity, strong mantle lithosphere that underlies the southern plateau^{10,11,25,26}. Significant northward underthrusting of the Indian lithosphere is consistent with recent modelling of the gravity field along this same transect²⁷ and teleseismic tomographic images²².

The low mean crustal velocity and low-to-normal Poisson's ratio observed here for the Lhasa terrane (and found by regional waveform modelling²³) has a consistent explanation in this inter-

pretation. We suggest that the underthrusting of India is accommodated by a series of ramp thrust faults that migrate progressively down through the crust, thus thickening the Lhasa terrane by tectonic underplating. This mechanism, in effect, transfers a series of thrust plates of Indian upper-crustal rocks to the underside of the Asian continental margin and accretes Indian affinity crust to the southern margin of Asia. This model predicts that the thickness of the Indian crust underlying the Lhasa terrane decreases northwards due to this mechanism of crustal transfer. The net effect is to construct a crust with predominantly upper-crustal rocks²⁸ that are low in velocity and low-to-normal in Poisson's ratio. This prediction is consistent with geometric constraints on the Main Himalaya thrust (MHT) from seismic reflection studies²⁹, recent geological mapping in the Lhasa terrane indicating the absence of Tertiary distributed shortening³⁰, and regional earthquake waveform modelling²³. Although crustal injection³¹ can also produce thickening without shortening, we prefer the tectonic underplating model because the low Poisson's ratio is more consistent with a compositional interpretation²⁴.

North of the Banggong suture, the crustal structure is quite different: the distinct high-velocity layer in the lower crust is absent, Poisson's ratio is higher, and the crustal thickness is less. Our models for the northern stations do show a smooth velocity increase in a crust–mantle transition zone, but lack the abrupt increase that produced the characteristic 'double-peak' Moho Sp arrival at the southern stations. The Qiangtang terrane, bounded by the Banggong and Jinsha River sutures, is characterized by intermediate values of Poisson's ratio (0.29) and thickness (65 km). The Songpan-Ganz terrane's crust between the Jinsha River suture and the Qaidam basin is thin (55 km) with a bulk Poisson's ratio of 0.33,

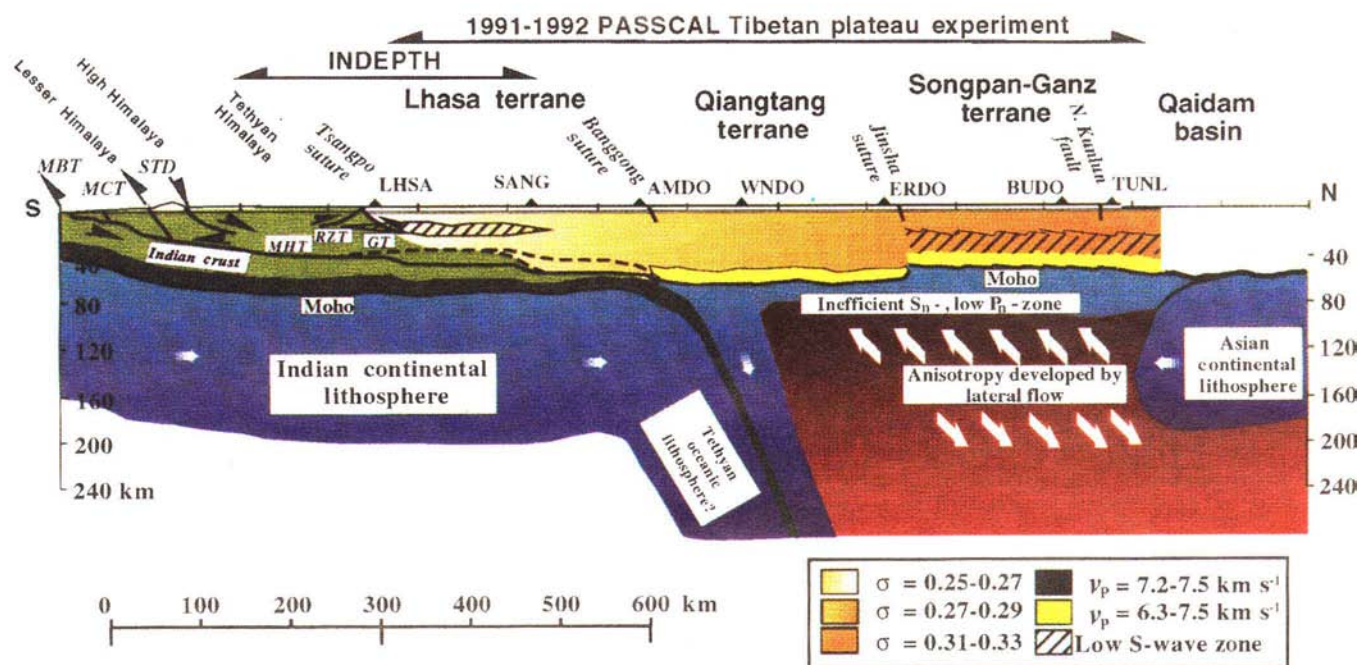


Figure 5 Interpretative north-south lithospheric cross-section in the central Tibetan plateau at about 92° to 93° E. Elements from the results of our experiment⁴, the two Sino-French experiments^{7,8} and the Sino-American INDEPTH project^{5,6} have been combined in this cross-section. The current study constrains the average P-wave velocity in the crust to be between 6.0 and 6.2 km s⁻¹ throughout the central Tibetan plateau. Crust of Indian affinity is shown in green to extend beneath the Lhasa terrane as expected from our tectonic underplating interpretation. Orange shading in the crust indicates observed variations in average crustal Poisson's ratio. The Poisson's ratios shown in the key are bulk estimates for all layers above the Moho, including green, dark grey, and yellow layers where appropriate. The sub-crustal properties are summarized from a variety of sources^{8-11,25,27,40,41}. Blue areas in the mantle indicate mantle lithosphere, although lid thicknesses are currently only constrained in south²³. Red areas indicate regions where teleseismic tomography²² and other data suggest that the

mantle is lower velocity relative to the southern plateau. The slowest region is at ~200 km depth beneath the Songpan-Ganz terrane⁹. The existence of steeply dipping Tethyan oceanic lithosphere beneath the central plateau is speculative, but would explain the deeply rooted northward-dipping velocity contrast in tomographic images²² at that location. This is also a natural explanation for the torque required by models of the gravity field²⁷. The presence of an advancing deeply rooted front would provide a mechanism to produce lateral flow and anisotropy in the mantle beneath the northern Tibetan plateau⁹. The Renbu-Zedong thrust system (RZT) and the Gangdese thrust (G) in southern Tibet were determined through geological mapping⁴⁰. The Main Boundary thrust (MBT), Main Central thrust (MCT), and Southern Tibet detachment (STD) are also based on geological mapping, but the Main Himalayan thrust (MHT) south of the Tsangpo suture is based on seismic reflection observations⁵. Properties of crust north of TUNL are not constrained by this study.

and is located above a low-velocity body at about 200 km depth⁸. The high Poisson's ratio of the Songpan-Ganz terrane is due to unusually low S-wave velocities in the lower crust suggesting prevalent partial melting, probably related to (or induced by) the high temperatures in the underlying mantle. Both the Qiangtang and Songpan-Ganz terranes are located above an anomalous mantle characterized by inefficient regional upper-mantle S-wave (S_n) propagation^{10,11}, low regional upper-mantle P-wave (P_n) velocity¹⁰ and strongly developed shear-wave anisotropy⁹.

The cross-section in Fig. 5 and the discussion above appears to imply a correlation between the crustal property variations and accreted crustal terranes. Inspection of the seismograms at USHU, located off the transect line but on strike with ERDO along the Jinsha suture, reveals a closer resemblance to WNDO than to ERDO or BUDO (Fig. 2). The closer correspondence of the change in crustal structure with the southern boundary of the inefficient S_n zone (that also diverges east of AMDO from the Banggong suture) rather than a terrane boundary casts some doubt on the apparent correlation of crustal structures and terranes. Nonetheless, results from the Sino-French experiment show evidence for an abrupt change in crustal properties at the Jinsha River suture^{8,32}. Resolution of this question awaits similar transect studies along the different portions of the Tibetan plateau.

Our determination of the thinner crust and very high Poisson's ratio in the crust, when combined with extensive evidence for high-

temperature upper mantle, is strongly suggestive of pervasive partial melting of the lower half of the crust in the northern plateau. However, owing to ambiguity associated with the effect of small melt fractions on bulk Poisson's ratio, it does not follow that the low-to-normal Poisson's ratio in the southern plateau is evidence against the presence of any partial melt within the crust of the Lhasa terrane. In fact, recent INDEPTH results have produced evidence of a thin layer (10–15 km) of partially molten material in the upper crust beneath the Yadong-Gulu rift in southern Tibet^{33–36}. This and other arguments^{31,37,38} led the INDEPTH team⁶ to speculate that the entire middle crust of the Tibetan plateau may be partially molten and play a major role in the mechanical behaviour of the plateau in the manner envisioned by Zhao and Morgan³¹. Our analysis resolves with confidence that the bulk Poisson's ratio is nearly normal in the southern plateau. If high Poisson's ratio is an appropriate proxy for melt accumulation, our results necessarily restrict such accumulations to thin layers (approximately 10–15 km). In this sense, our observations are compatible with the basic observations^{33–36} of the INDEPTH experiment. Additionally, thermal modelling³⁹ suggests that our preferred mechanism of crustal thickening by tectonic underplating could, in theory, generate crustal melting. However, reconciliation of the inference of a widespread, thick partially molten middle crust in the southern plateau⁶ with our observations of low bulk Poisson's ratio will require further analysis and perhaps greater spatial coverage of observations within the plateau. For

example, a detailed analysis of the relative roles of the rate of crustal thickening (which lowers Poisson's ratio in our interpretation) and the development of crustal melting (which would raise Poisson's ratio) is needed. Further, understanding the role that both of these phenomena may play in the modifying the mechanical behaviour of the crust is crucial.

In the north, we cannot constrain exactly the depth at which the crustal melting exists, but we can say with confidence that Poisson's ratios in the range of 0.34–0.36 must exist throughout a depth range of more than 30 km. We have correlated the top of a known crustal low-velocity zone²¹ with the onset of this high-Poisson's-ratio region of the crust. We infer that the middle and lower crust of the northern plateau must be mechanically weak and capable of flow, and suggest that the influx of heat from the anomalous upper mantle is a major influence on the development of this anomalous crustal structure. At present, it is not possible to conclude that heat from the upper mantle is the only factor in the development of this zone. Crustal thickening has certainly occurred in the northern plateau and, thus, it is possible that crustal heating³⁹ took place before the influx of heat from the mantle. In terms of the evolution of the plateau, it is important to try to resolve this ambiguity. Further detailed study of the transitional region near 32°N may be crucial in resolving the relative importance of the various factors for the structure and mechanical evolution of the Tibetan plateau.

The nature of the mantle underlying the northern Tibetan plateau remains somewhat paradoxical. Many seismic indicators point to the presence of hot, partially melted mantle just below the crust, that is, asthenosphere, although interpretations of P_n tomographic images⁴⁰ and geochemical data⁴¹ suggest that at least a thin mantle lithospheric lid exists in the north. Additionally, the strong shear-wave anisotropy beneath in the northern plateau has been interpreted as an indicator of highly strained lithosphere⁹. In such a model, the Asian margin lithosphere has been shortened and heated, either by internal strain or by increased heat from below, inducing some melting in the lid⁴¹. An alternative idea is that the pre-existing Asian mantle was thin and weak and is being scraped off and recycled into the mantle by the underthrusting Indian lithosphere. The remaining asthenosphere is then squeezed between the advancing Indian front and the backstop of the older Asian continental lithosphere (Fig. 5). The squeezing induces a tightly constrained lateral flow of the asthenosphere that develops a strong anisotropic fabric⁴². The resolution of these conflicting upper-mantle models awaits a better understanding of the rheology and mineral physics of mantle materials.

There are few geological constraints on the timing and mode of uplift in the northern Tibetan plateau. We suggest that the anomalous mantle and crustal properties we report here are geologically young features triggered by progressive heating, squeezing and lateral flow of the Asian crust and mantle between two strong converging lithospheres. The correlation between areas of thin crust and an extremely high-Poisson's-ratio lower crust is strongly suggestive of a pervasively melted, near-fluid, lower crust flowing to maintain a uniform topography. A thick, weak crust would flow to maintain uniform elevations while the Moho would either flatten in the absence of mantle variations or take on topography in the presence of such variations⁴³. We note that lower-crustal flow is not the same as crustal extrusion and, in fact, the two processes are probably mutually exclusive. Crustal flow is a high-temperature process that implies a decoupling of lower-crustal motions from the upper crust (for which there is evidence in the present gravity field in the plateau³⁸). Crustal extrusion is a low temperature process that implies movement of discrete crustal blocks between lithosphere-penetrating faults⁴⁴. In this context, our interpretation of crustal flow in the northern plateau is consistent with models that have crustal extrusion in this area occurring early in the collision process⁴⁵. The heating and eventual melting of the crust over time would provide a mechanism for the transition from block extrusion

to crustal flow as a response to the continuing collision.

The important general observation is that the convergence of India and Asia has been accommodated in very different ways in the southern and northern plateau. In the south, at least 350 km of convergence⁴⁶ has been accommodated largely by underthrusting along crustal-scale ramp-thrust faults (for example, the Main Himalayan thrust⁵). Although thrusting occurred in the north early in the collision⁴⁵, the northern Tibetan plateau crust is now melted (and perhaps flowing) owing to heating by the high-temperature, highly strained mantle beneath it. This difference is also manifest in different modes of isostatic support of the high uniform elevations. The south is supported principally by thick, low-density crust whereas the north is supported by hot, low-density upper mantle. The relatively uniform elevations (of approximately 5 km) are unlikely to be a coincidence of these two mechanisms working independently. It is likely that the 5-km-average elevation represents some maximum potential-energy level and that after achieving the maximum elevation, further thickening is deflected to gravitational collapse⁴⁷, northward migration of the deformation front⁴⁷ or lower-crustal flow⁴³. □

Received 15 July 1996; accepted 11 March 1997.

- Fielding, E., Isacks, B., Barazangi, M. & Duncan, C. How flat is Tibet? *Geology* **22**, 163–167 (1994).
- Brandon, C. & Romanowicz, B. A "no-lid" zone in the central Chang-Thing Platform of Tibet: Evidence from pure path velocity measurements of long period Rayleigh waves. *J. Geophys. Res.* **91**, 6547–6564 (1986).
- Molnar, P. A review of geophysical constraints on the deep structure of the Tibetan Plateau, the Himalaya, and the Karakoram, and their tectonic implications. *Phil. Trans. R. Soc. Lond. A* **326**, 33–88 (1988).
- Owens, T. J., Randall, G. F., Wu, F. T. & Zeng, R. PASSCAL instrument performance during the Tibetan Plateau passive seismic experiment. *Bull. Seismol. Soc. Am.* **83**, 1959–1970 (1993).
- Zhao, W. *et al.* Deep seismic reflection evidence for continental underthrusting beneath southern Tibet. *Nature* **366**, 557–559 (1993).
- Nelson, D. *et al.* Partially molten middle crust beneath southern Tibet: Synthesis of Project INDEPTH results. *Science* **274**, 1684–1687 (1996).
- Hirn, A. *et al.* Seismic anisotropy as an indicator of mantle flow beneath the Himalayas and Tibet. *Nature* **375**, 571–574 (1995).
- Wittlinger, G. *et al.* Seismic tomography of northern Tibet and Kunlun: Evidence for crustal blocks and mantle velocity contrasts. *Earth Planet. Sci. Lett.* **139**, 263–279 (1996).
- McNamara, D. E., Owens, T. J., Silver, P. G. & Wu, F. T. Shear wave anisotropy beneath the Tibetan Plateau. *J. Geophys. Res.* **99**, 13655–13665 (1994).
- McNamara, D. E., Owens, T. J. & Walter, W. R. Observations of regional phase propagation across the Tibetan Plateau. *J. Geophys. Res.* **100**, 22215–22229 (1995).
- Ni, J. & Barazangi, M. High-frequency seismic wave propagation beneath the Indian shield, Himalayan arc, Tibetan Plateau and surrounding regions: High uppermost velocities and efficient propagation beneath Tibet. *Geophys. J. R. Astron. Soc.* **72**, 665–689 (1983).
- Jordan, T. J. & Frazer, L. N. Crustal and upper mantle structure from Sp phases. *J. Geophys. Res.* **80**, 1504–1518 (1975).
- Zandt, G. & Randall, G. E. Observations of shear-coupled P-waves. *Geophys. Res. Lett.* **12**, 565–568 (1985).
- Baag, C. E. & Langston, C. A. Diffracted Sp generated under the Australian Shield. *Geophys. J. R. Astron. Soc.* **80**, 363–386 (1985).
- Langston, C. A. The Ssmp phase in regional wave propagation. *Bull. Seismol. Soc. Am.* **86**, 133–143 (1996).
- Bath, M. & Stefansson, R. S-P conversion at the base of the crust. *Ann. Geofis.* **19**, 119–130 (1966).
- Randall, G. E. Efficient calculation of complete differential seismograms for laterally homogeneous earth models. *Geophys. J. Int.* **118**, 245–254 (1994).
- Zandt, G. & Ammon, C. J. Continental crust composition constrained by measurements of crustal Poisson's ratio. *Nature* **374**, 152–154 (1995).
- Zhao, L.-S., Sen, M. K., Soffa, P. & Frohlich, C. Application of very fast simulated annealing of the determination of the crustal structure beneath Tibet. *Geophys. J. Int.* **125**, 355–370 (1996).
- Zhu, L., Zeng, R., Wu, F., Owens, T. & Randall, G. Preliminary study of crust-upper mantle structure of the Tibetan Plateau by using broadband teleseismic body waveforms. *Acta Seismol. Sinica* **6**, 305–316 (1993).
- Zhu, L., Owens, T. J. & Randall, G. E. Lateral variation in crustal structure of the northern Tibetan Plateau inferred from teleseismic receiver functions. *Bull. Seismol. Soc. Am.* **85**, 1531–1540 (1995).
- Vandecar, J. & McNamara, D. E. 3-D upper-mantle velocity structure beneath the Tibetan Plateau (abstr.). *Fos 77*, F694 (1996).
- Rodgers, A. J. & Schwartz, S. Y. Low crustal velocities and mantle lithospheric variations in southern Tibet from regional Pnl waveforms. *Geophys. Res. Lett.* **24**, 9–12 (1997).
- Christiansen, N. Poisson's ratio and crustal seismology. *J. Geophys. Res.* **102**, 3139–3156 (1996).
- Beghoul, N., Barazangi, M. & Isacks, B. Lithospheric structure of Tibet and western North America: Mechanisms of uplift and a comparative study. *J. Geophys. Res.* **98**, 1997–2016 (1993).
- Holt, W. E. & Wallace, T. C. Crustal thickness and upper mantle velocities in the Tibetan Plateau region from the inversion of regional Pnl waveforms: Evidence for a thick upper mantle lid beneath southern Tibet. *J. Geophys. Res.* **95**, 12499–12525 (1990).
- Jin, Y., McNutt, M. K. & Zhu, Y. Mapping the descent of Indian and Eurasian plates beneath the Tibetan Plateau from gravity anomalies. *J. Geophys. Res.* **101**, 11275–11290 (1996).
- Butler, R. W. H. Thrust tectonics, deep structure, and crustal subduction in the Alps and Himalayas. *J. Geol. Soc. Lond.* **143**, 857–873 (1986).
- Makovskiy, Y. *et al.* Structural elements of the southern Tethyan Himalaya crust from wide-angle seismic data. *Tectonics* **15**, 997–1005 (1996).
- Yin, A., Murphy, M. A. & Harrison, T. M. Significant crustal shortening in the Lhasa block (southern Tibet) predates Indo-Asian collision. *Geol. Soc. Am. Abstr. Program* **27**, A-335 (1995).
- Zhao, W.-L. & Morgan, W. J. Injection of Indian crust into Tibetan lower crust: A two-dimensional

- finite element model study. *Tectonics* **6**, 489–504 (1987).
32. Herquel, G., Wittlinger, G. & Guilbert, J. Anisotropy and crustal thickness of Northern-Tibet. New constraints for tectonic modeling. *Geophys. Res. Lett.* **22**, 1925–1928 (1995).
 33. Brown, L. D. *et al.* Bright spots, structure and magmatism in southern Tibet from INDEPTH seismic reflection profiling. *Science* **274**, 1688–1690 (1996).
 34. Makovsky, Y. *et al.* INDEPTH Wide-angle reflection observation of P-wave to S-wave conversion from crustal bright spots in Tibet. *Science* **274**, 1690–1691 (1996).
 35. Kind, R. *et al.* Evidence from earthquake data for a partially molten crustal layer in southern Tibet. *Science* **274**, 1692–1694 (1996).
 36. Chen, L. *et al.* Electrically conductive crust in southern Tibet from INDEPTH magnetotelluric surveying. *Science* **274**, 1694–1696 (1996).
 37. Francheteau, J. *et al.* High heat flow in southern Tibet. *Nature* **307**, 32–36 (1984).
 38. Jin, Y., McNutt, M. K. & Zhu, Y. Evidence from gravity and topography for folding of Tibet. *Nature* **371**, 669–674 (1994).
 39. England, P. C. & Thompson, A. in *Collision Tectonics* (eds Coward, M. P. & Ries, A.C.) 83–94 (Spec. Publ. 19, Geol. Soc. Lond., 1986).
 40. McNamara, D. E., Walter, W. R., Owens, T. J. & Ammon, C. J. Upper mantle velocity structure beneath the Tibetan Plateau from Pn travel time tomography. *J. Geophys. Res.* **102**, 493–505 (1997).
 41. Turner, S. *et al.* Timing of the Tibetan uplift constrained by analysis of volcanic rocks. *Nature* **364**, 50–54 (1993).
 42. Zhang, S. & Karato, S. Lattice preferred orientation of olivine aggregates deformed in simple shear. *Nature* **375**, 774–777 (1995).
 43. Bird, P. Lateral extrusion of lower crust from under high topography, in the isostatic limit. *J. Geophys. Res.* **96**, 10275–10286 (1991).
 44. Tappinier, P., Peltzer, G. & Armijo, R. in *Collision Tectonics* (eds Coward, M. P. & Ries, A.C.) 115–157 (Spec. Publ. 19, Geol. Soc. Lond., 1986).
 45. Leeder, M. R., Smith, A. B. & Yin, J. Sedimentology, paleoecology and paleoenvironmental evolution of the 1985 Lhasa to Golmud Geotraverse. *Phil. Trans. R. Soc. Lond. A* **327**, 107–143 (1988).
 46. Harrison, T. M., Copeland, P., Kidd, W. S. F. & Yin, A. Raising Tibet. *Science* **255**, 1663–1670 (1992).
 47. Molnar, P. & Lyon-Caen, H. *Some Simple Physical Aspects of Mountain Building* 179–207 (Spec. Pap. 218, Geol. Soc. Am., Boulder, CO, 1988).
 48. Yin, A. *et al.* Tertiary structural evolution of the Gangdese thrust system in southeastern Tibet. *J. Geophys. Res.* **99**, 18175–18201 (1994).

Acknowledgements. We thank the American and Chinese scientists whose work in the field resulted in this data set, especially R. Zeng of the Institute of Geophysics, State Seismological Bureau, China and F. Wu, SUNY-Binghamton; D. McNamara for contributing to numerous discussions and providing Pg velocity estimates; A. Yin, K. Furlong, D. Nelson and E. Sandvol for providing comments. Figure 1 was prepared using the GMT software package by P. Wessel and W. Smith. This work was supported by the US NSF and AFOSR. T.O. thanks the University of Washington for space and support during this work.

Correspondence should be addressed to T.J.O. (e-mail: owens@sc.edu).

A complex containing N-CoR, mSin3 and histone deacetylase mediates transcriptional repression

Thorsten Heinzel*, Robert M. Lavinsky*†, Tina-Marie Mullen‡, Mats Söderström§, Carol D. Laherty||, Joseph Torchia*, Wen-Ming Yang†, Gyan Brard§, Sally D. Ngo§, James R. Davie#, Edward Seto†, Robert N. Eisenman||, David W. Rose‡, Christopher K. Glass§ & Michael G. Rosenfeld*

* Howard Hughes Medical Institute, † UCSD Graduate Program in Biology, ‡ Whittier Diabetes Program, and § Cellular and Molecular Medicine, Department and School of Medicine, University of California, San Diego, La Jolla, California 92093-0648, USA

|| Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98104, USA

§ H. Lee Moffitt Cancer Center and Research Institute, Department of Medical Microbiology and Immunology, University of South Florida, Tampa, Florida 33612, USA

Department of Biochemistry and Molecular Biology, University of Manitoba, Winnipeg, Manitoba, R3E 0W3, Canada

Transcriptional repression by nuclear receptors has been correlated to binding of the putative co-repressor, N-CoR. A complex has been identified that contains N-CoR, the Mad presumptive co-repressor mSin3, and the histone deacetylase mRPD3, and which is required for both nuclear receptor- and Mad-dependent repression, but not for repression by transcription factors of the ets-domain family. These data predict that the ligand-induced switch of heterodimeric nuclear receptors from repressor to activator functions involves the exchange of complexes containing histone deacetylases with those that have histone acetylase activity.

Thyroid-hormone and retinoic-acid receptors belong to the nuclear receptor family of ligand-dependent transcription factors that control development and homeostasis as a consequence of both positive and negative control of gene expression^{1,2}. These receptors, as well as the viral oncogene *erbA* and *RevErb*, actively repress transcription of target genes in the absence of ligand^{1,2}, although maximal efficiency occurs only in the presence of chromatin³. The ligand-binding domains of the thyroid-hormone and retinoic-acid receptor transfer active repression function to the GAL4 DNA-binding domain (amino acids 1–147)^{4,5}. The repression activity is dependent on a conserved region in the amino-terminal part of the ligand-binding domain^{6,7}, termed the CoR-box⁸. These observations led to the biochemical identification and cloning of a protein of relative molecular mass 270,000 (*M_r* 270K), referred to as N-CoR, which binds to the ligand-binding domain of thyroid-hormone and retinoic-acid receptors^{8,9}, or to comparable regions of *RevErb*¹⁰, by means of adjacent carboxy-terminal binding domains. Receptor

mutations that abolished N-CoR binding *in vitro* but had no effect on receptor DNA binding or ligand-dependent transactivation were found to inhibit repressor function. These data, and the presence of transferable repressor domains, led to the proposal that N-CoR might mediate ligand-independent transcriptional repression⁸.

Analogous to repression functions of unliganded nuclear receptors, the bHLHZ proteins of the Mad family act as transcriptional repressors on heterodimerization with Max, and are believed to be involved in the induction of terminal differentiation in a wide range of cell types^{11–13}. The repressor function of Mad has been correlated to the recruitment of mammalian Sin3 orthologues (mSin3A and mSin3B) to a highly conserved putative amphipathic helix at its N terminus, referred to as the Sin3-interaction domain (Mad-SID)^{14–16}. Yeast Sin3 was initially identified in genetic screens for suppressor mutations of a *Swi5* defect in yeast, suggesting that it is involved in chromatin rearrangement^{17–20}. It was subsequently linked by epistasis in yeast to the global transcriptional regulator

Here we report the identification of a cellular N-CoR/mSin3/mRPD3 co-repressor complex, and show that N-CoR, mSin3 and the mRPD3 histone deacetylase are essential components for both nuclear receptors and Mad-mediated repression. We also demonstrate that this co-repressor complex is specific for certain classes of transcriptional repressors.

To investigate the potential of a mechanistic relationship between nuclear receptor and Mad-dependent repression, we generated antibodies against different regions of N-CoR; these were used to analyse potential nuclear complexes. Cellular extracts from MCF-7 cells were immunoprecipitated using affinity-purified anti-mSin3A/B or anti-mRPD3 IgGs. After washing and SDS-polyacrylamide gel electrophoresis, western blot analysis was performed with anti-N-CoR IgG. This analysis revealed that both mSin3 proteins and mRPD3 exist in a cellular complex containing N-CoR (Fig. 1a). To investigate whether the interactions of mRPD3 and mSin3 with N-CoR were direct, baculovirus-expressed N-CoR was affinity-purified and used in immunoprecipitations with ³⁵S-methionine-labelled mSin3B or mRPD3. N-CoR was found to bind directly mSin3B (and mSin3A; data not shown), but not mRPD3 (Fig. 1b). Although several experimental approaches failed to detect direct interaction, anti-N-CoR antibodies immunoprecipitated substan-

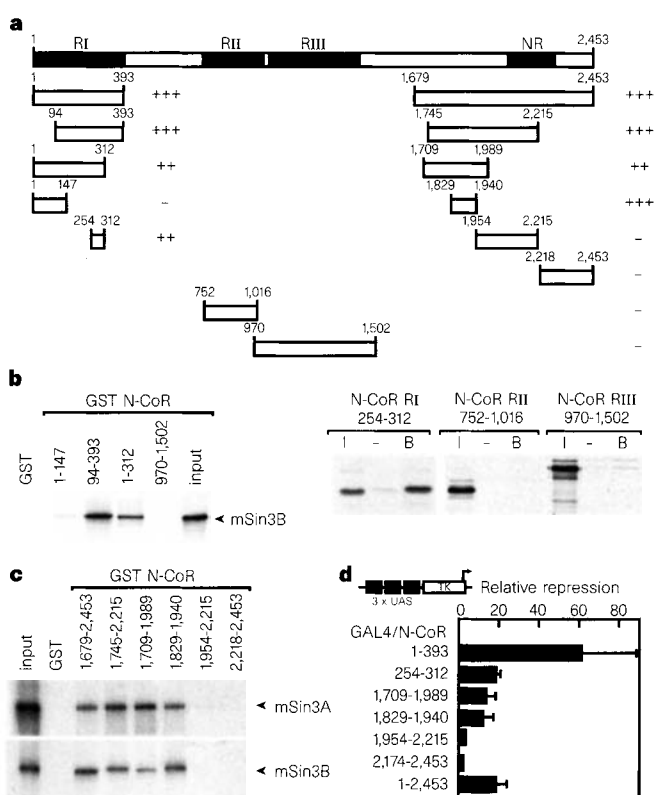


Figure 2 Regions of N-CoR involved in interaction with mSin3. **a**, The indicated GST fusions of N-CoR were incubated with ³⁵S-labelled mSin3A or mSin3B. Specifically bound mSin3 was detected by autoradiography and the relative strength of interaction was scored from – (background level) to + + + (strong). **b**, Interaction of GST fusions of N-CoR, or GST protein as a control, with mSin3B (left). ³⁵S-labelled NCoR amino acids 254–312 and fragments encompassing RII or RIII were immunoprecipitated using antibodies against mSin3 (right), or in the absence of mSin3-specific antibodies (–). **c**, Interaction of GST fusions of C-terminal regions of N-CoR with mSin3. **d**, The indicated regions of N-CoR were fused to GAL4 (amino acids 1–147) and tested for repression of a 3x-UAS-TK-luciferase reporter.

tial histone deacetylase activity from HeLa cells (Fig. 1c). Together, these data suggest the existence of nuclear complexes containing N-CoR, both mSin3 proteins and mRPD3.

To map the specific domain in N-CoR that mediates interactions with mSin3, a series of glutathione S-transferase (GST) fusion proteins representing overlapping portions of N-CoR were expressed in bacteria, purified and used to bind ³⁵S-labelled mSin3A or mSin3B. Our results revealed the presence of two interaction domains at amino acids 1–393 (repressor domain I) and 1,745–2,215 (Fig. 2a,b). Immunoprecipitations demonstrated that repressor domains II and III in N-CoR did not interact with mSin3 (Fig. 2b). The N-terminal interaction domain of N-CoR was more precisely defined to a sequence encompassing amino acids 254–312 (Fig. 2b), whereas C-terminal amino acids 1,829–1,940 were sufficient for the interaction with mSin3A/B (Fig. 2c).

To investigate the ability of the minimal domains of N-CoR that interact with mSin3 to transfer repression, cotransfection experiments were performed. GAL4 (1–147)–N-CoR fusion proteins were tested for their ability to repress a transcription unit under the control of the complete human simian virus thymidine kinase (TK) promoter with three GAL4-binding (UAS) sites. Indeed, GAL4 fusions of both mSin3-interaction domains in N-CoR exhibited significant repression (baseline was taken to be 260,000 relative light units; Fig. 2d).

To determine the precise regions in mSin3A and mSin3B that interact with N-CoR, a series of ³⁵S-labelled fragments of each mSin3 were generated (mSin3A, PAH1 (amino acids 1–205), PAH1 → 2 (1–479), PAH1 → 3 (1–680), PAH1 → 4 (1–1,015), Fig. 3a; mSin3B, PAH1 (1–173), PAH1 → 2 (1–293), PAH1 → 3 (1–485), PAH3 (275–499); Fig. 3b). This analysis showed that amino acids 275–499 of mSin3B, encompassing PAH3 and part of the linker between PAH3 and PAH4, were sufficient to mediate interaction with the N terminus of N-CoR. In contrast, none of the mSin3 fragments in which PAH3 was deleted interacted with GST–N-CoR amino acids 1–312. Fragments of mSin3 containing PAH1 (amino acids 1–205 of mSin3A, and 1–173 of mSin3B) were sufficient for interaction with the C-terminal domain of N-CoR, however (Fig. 3b). As the N-termini of mSin3A and mSin3B are highly divergent, our data suggest that the interaction with the N-CoR C terminus is mediated by the conserved PAH1 domain.

Roles of N-CoR, mSin3 and mRPD3 in repression

To evaluate directly the potential roles of components of the N-CoR/mSin3/mRPD3 complex in nuclear receptor and Mad repression, we used the single-cell microinjection assay²⁷. Expression of a GAL4/thyroid-hormone receptor fusion protein (TRC' amino acids 165–465) resulted in marked repression of a *lacZ* reporter gene under the control of a UAS-dependent TK promoter (Fig. 4). In this assay, microinjection of anti-N-CoR IgG abolished the repression imparted by the thyroid-hormone receptor C terminus, demonstrating that N-CoR serves as an essential nuclear receptor co-repressor. Further, anti-mSin3 IgG strongly relieved repression by the GAL4/thyroid-hormone receptor. Finally, the effect of anti-mRPD3 IgG on nuclear receptors led to a marked, albeit incomplete, reversal of repression (Fig. 4a, b). Given that the mSin3 antibody nearly completed reversal of inhibition, it is likely that other deacetylases in the mSin3-containing complex also mediate repression events, consistent with our finding that the deacetylase HDAC1 can co-immunoprecipitate with mSin3B *in vitro* (data not shown).

The effects of mSin3 and mRPD3 on thyroid-hormone receptor led us to determine whether they were directly required for N-CoR-mediated repression. Microinjection of either anti-mSin3A/B or anti-mRPD3 IgG completely blocked repression by the GAL4–N-CoR fusion protein, indicating that mRPD3 and mSin3 are likely to exert their effects on the thyroid-hormone receptor through N-CoR (Fig. 4b). To ensure that the effects on thyroid-hormone receptor-mediated repression were specific, rather than reflecting a general

increase in transcription by interference with endogenous histone deacetylases, we tested all IgGs on a series of promoters of different strengths. None of these antibodies exerted any significant effect on transcription from these promoters (Fig. 4c).

Our data strongly suggest that recruitment of histone deacetylases plays an important role in both nuclear-receptor and Mad-mediated repression. To further test this hypothesis we analysed the effects of the irreversible histone deacetylase inhibitor, trapoxin²⁵, on GAL4 fusion proteins (GAL4–TRC', GAL4–N-CoR, GAL4–Mad1) and on GAL4 alone, using a promoter containing three GAL4-binding sites. Although a small increase in transcription of the GAL4 control was observed, the GAL4 fusions of the thyroid hormone receptor, N-CoR and Mad1 consistently showed at least a three- to fourfold decrease in repression (Fig. 4d).

The role of the co-repressor complex in Mad-mediated repression was evaluated further using the microinjection assay (Fig. 5). Consistent with the well-established interaction between Mad1 and mSin3 (refs 14, 16), combined anti-mSin3A/B IgGs virtually abolished Mad1-dependent repression. Either anti-mSin3A- or anti-mSin3B-specific IgG alone resulted in only partial reversal of repression, suggesting that both mSin3A and mSin3B contribute to Mad-dependent inhibitory events (Fig. 5). Further, anti-mRPD3 IgG effectively relieved Mad-dependent repression, consistent with epistatic data from yeast. Because N-CoR does not interact directly with Mad1 (data not shown) or mRPD3, and requires mSin3 for its own repression of thyroid-hormone receptor, we sought to determine whether N-CoR was in turn required for repression by Mad.

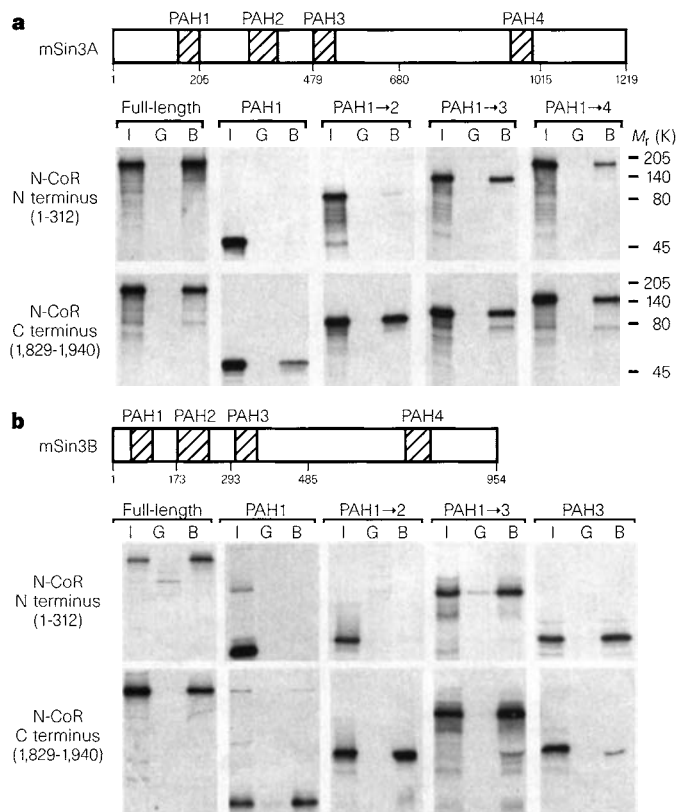


Figure 3 Domains within mSin3A and mSin3B required for N-CoR interactions. **a**, Schematic representation of mSin3A indicating the locations of PAH domains 1–4. Interaction of the indicated ³⁵S-labelled myc-tagged mSin3A deletion mutants or full-length protein with regions of N-CoR (amino acids 1–312 or 1,829–1,940) was demonstrated by incubation with GST fusion proteins (B) or GST alone (G) bound to glutathione agarose; 10% of input protein is shown (I). **b**, Schematic representation of mSin3B. Interaction of mSin3B deletion mutants or full-length protein with GST fusion proteins of N-CoR was performed as described above.

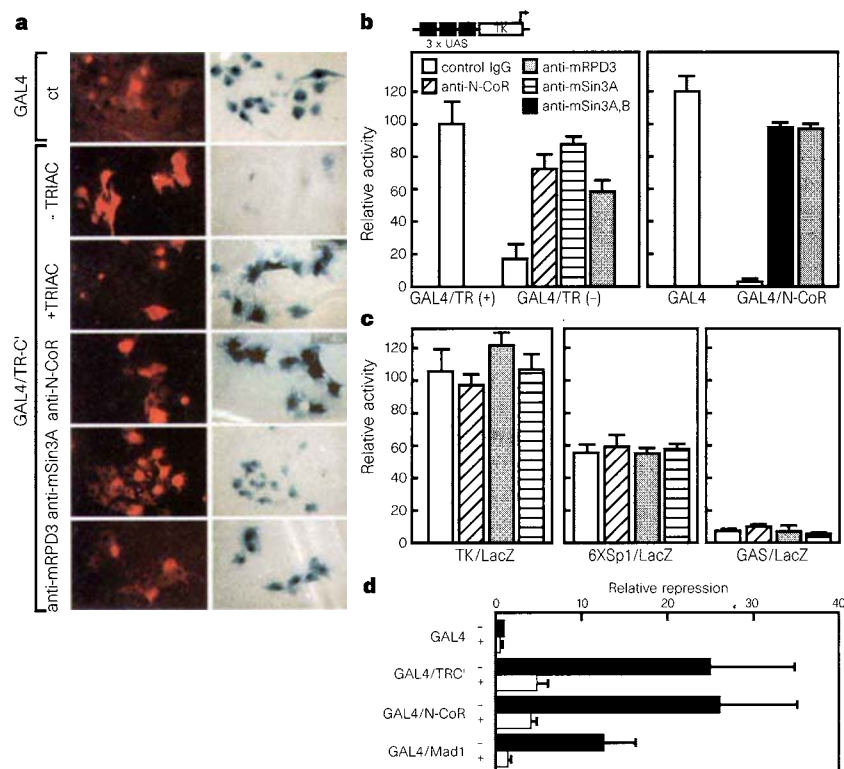


Figure 4 Microinjection of antibodies against N-CoR, mSin3 and mRPD3 abrogates transcriptional repression by N-CoR and TR. A 3x-UAS-TK-lacZ reporter was injected into the nuclei of Rat-1 cells and quantified by the percentage of injected cells (by rhodamine staining) that also turned blue (X-gal staining). **a**, Rhodamine (left) and X-gal staining (right) for the indicated constructs and antibodies. **b**, The percentage of X-gal-stained injected cells was normalized to the baseline ($79 \pm 10\%$) and given as relative activity. **c**, Antibodies were co-injected with control promoters. **d**, Mad1-SID (GAL4/Mad1), the ligand-binding domain of T₃R-β (GAL4/TRC'), or N-CoR were fused to GAL4 (1–147) and cotransfected with the 3x-UAS-TK-luciferase reporter to determine transcriptional repression in the absence or presence of 10^{-8} M trapoxin. Repression is given relative to GAL4.

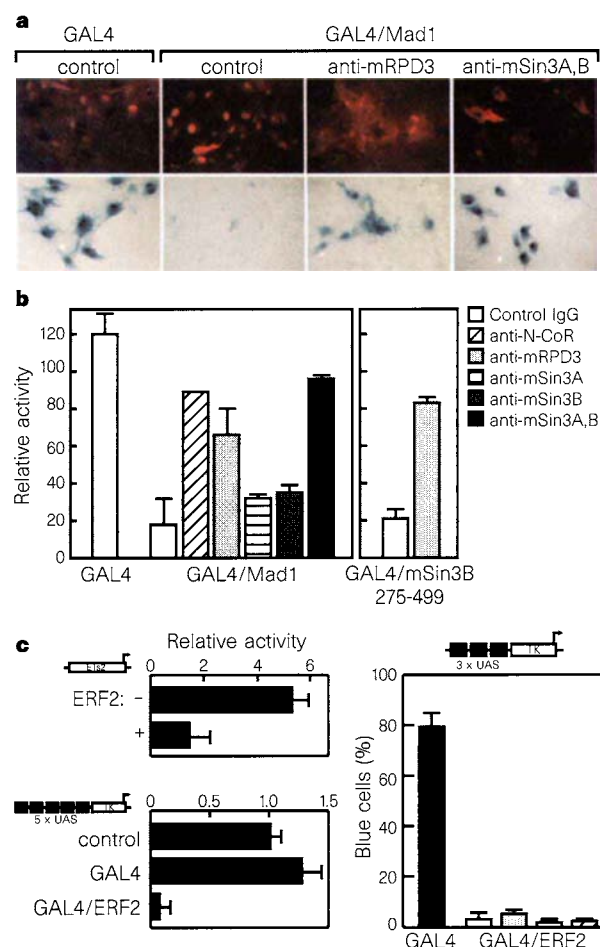


Figure 5 mSin3 and mRPD3, as well as N-CoR, are required for repression by Mad. Microinjections into Rat-1 and 293 cells (for anti-N-CoR on GAL4/Mad1) are shown. **a**, Rhodamine and X-gal staining for the indicated constructs and antibodies. **b**, Relative activities for the indicated expression plasmids and IgGs are normalized to the baseline obtained with the reporter plasmid alone. **c**, Repression by ERF2 and an GAL4-ERF2 (amino acids 130–531) fusion was analysed by transfection in HeLa cells with the indicated reporter plasmids. GAL4-ERF2 was microinjected with the 3x-UAS-TK-lacZ reporter and the indicated antibodies.

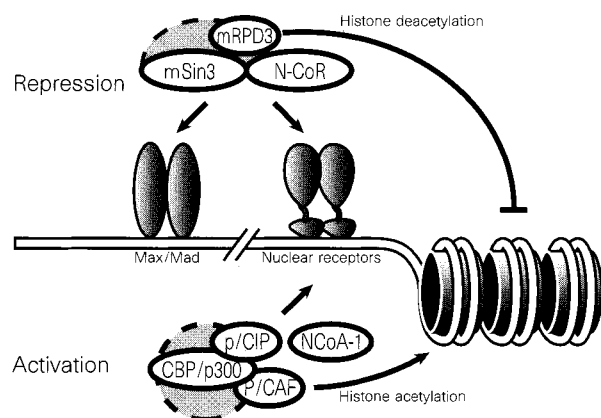


Figure 6 A model for the action of co-repressor and co-activator complexes mediating the transcriptional effects of nuclear receptors and Mad/Max heterodimers. Transcriptional repression requires a complex containing N-CoR, mSin3 and mRPD3 for local histone deacetylation. Upon binding of activating ligands, the co-repressor complex dissociates from nuclear receptors and is replaced by a co-activator complex containing CBP (ref. 46), a factor termed p/CIP (J.T. *et al.*, manuscript submitted), N-CoA1/SRC-1 (ref. 49) and P/CAF⁵⁰, which acetylates histones^{44,45} and facilitates the entry of core transcription factors.

Surprisingly, anti-N-CoR (amino acids 1–312) IgG abolished Mad-dependent repression, indicating that repression by Mad protein is dependent on at least three distinct components: N-CoR, mSin3 and histone deacetylases. To address whether this N-CoR effect is specific, rather than merely a reflection of depletion of an N-CoR-containing complex, an antibody against the nuclear receptor interaction domain in the C terminus of N-CoR (amino acids 2,239–2,453) was microinjected. In contrast to the N-terminally directed anti-N-CoR IgG (amino acids 1–312), the C-terminal IgG failed to inhibit Mad-dependent repression, although it effectively blocked thyroid-hormone receptor-mediated repression. Thus the antibodies against distinct regions of N-CoR selectively block distinct pathways mediated by N-CoR; these data exclude the possibility that the anti-N-CoR IgG acted by removing the entire complex.

To confirm that the effects of microinjected antibodies directed against N-CoR, mSin3 and mRPD3 are specific for reversing repression by nuclear receptors and Mad, rather than being the consequence of a general derepression, we investigated whether these antibodies also functioned to decrease active repression mediated by another class of repressor. Several members of the ets-domain family of transcription factors have been identified that act as repressors^{28,29}. One such transcription factor is ERF2, which is highly related to the ets repressor factor ERF²⁸ (G.B. and C.K.G., unpublished data). ERF2 exhibits a similar DNA recognition profile to ERF and, when expressed in HeLa cells, represses transcription from the ets2 promoter (Fig. 5c; G.B. and C.K.G., unpublished data). Transfer of the C-terminal domain of ERF2 to GAL4 results in repression in transfection (Fig. 5c) and microinjection assays. However, in contrast to the results obtained for Mad and nuclear receptors, microinjection of antibodies directed against N-CoR, mSin3 or mRPD3 had no significant effect on repression by ERF2. These observations suggest that there are several independent mechanisms for transcriptional repression.

Discussion

N-CoR and a related factor SMRT³⁰ were initially discovered through their ability to bind to unliganded nuclear receptors, whereas mSin3A/B were identified as putative co-repressors binding directly to a domain of Mad1 or Mxi^{14,15} required for repression. Here we demonstrate that both N-CoR and mSin3 are essential components of a co-repressor complex that is required for transcriptional repression in at least two classes of DNA-binding proteins, and that N-CoR acts as a co-repressor of nuclear receptors in the intact cell. Our data also suggest that N-CoR and mSin3 are the predominant functionally relevant factors that interact directly with the repression domains of nuclear receptors and Mad, respectively.

However, because no direct interactions could be detected between either N-CoR and mRPD3, or mSin3A/B and mRPD3 (ref. 51), it is likely that the co-repressor complex contains additional components. As predicted, based on the identification of

yeast Sin3/RPD1 and its functional epistasis to the genetically linked histone deacetylase RPD3 (ref. 21), there is a complex in mammalian cells that requires mSin3 and the mRPD3 histone deacetylase for repression (see also ref. 51). Recent data show that there are at least two complexes containing RPD3-like histone deacetylases in *Saccharomyces cerevisiae*, and that only one of these complexes is inhibited efficiently by deacetylase inhibitors such as trapoxin³¹.

Repression, a critical control mechanism in all aspects of biological processes, has been categorized as either active or passive, and long range or short range^{32–37}. Although many mechanisms have been proposed for repression by transcription factors, including interactions with the basal transcription apparatus, there is growing evidence that, at least in part, repression events may result from regulated alterations of chromatin structure by multi-protein complexes^{32–39}. Our model is consistent with studies in *Xenopus* showing that TR/RXR heterodimers use the chromatin environment for maximal repression⁴⁰. Components of the co-repressor complex include N-CoR, as well as mSin3 and mRPD3, in the cases of nuclear receptors and Mad (Fig. 6). These factors exert effects that are required, but are not individually sufficient, for the observed repression. In this regard, it is intriguing that splice variants of N-CoR that delete the exon(s) encoding the C-terminal mSin3-interaction domain have been identified⁴¹ (A. Hörlein and M.G.R., unpublished data) that could generate an N-CoR variant decoupled from deacetylase recruitment. However, our studies with the ets repressor ERF2 suggest that recruitment of this co-repressor complex is not likely to be universal. Additionally, the TUP1–SSN6 system, genetically linked to dominant active repression by MAT- α 2 (ref. 42), is not dependent on yeast Sin3. TUP1 has been reported to bind to histones H3 and H4 even when acetylated, and to cause altered nucleosome localization in DNA that is not normally transcribed³⁷; these observations suggest that its actions are not explained solely by interference with the basal transcriptional machinery. Furthermore, remodelling of chromatin structure and nucleosome positioning by targeting histone acetylation has been linked to gene activation^{38–40} and to establishing stable states of differential gene activity during *Xenopus* gastrulation⁴³. Indeed, CBP, which has intrinsic histone acetylation activity^{44,45}, is a component of a co-integrator complex⁴⁶ that includes a factor referred to as p/CIP (J.T., unpublished) and the associated histone acetylase P/CAF⁵⁰ (Fig. 6).

The existence of a common co-repressor complex, individual components of which directly bind to nuclear receptors and to Mad, could coordinate two distinct developmental and signalling pathways. We have shown that the ligand-induced switch from transcriptional repression to activation by nuclear receptors reflects the exchange of a histone deacetylase-containing co-repressor complex for an activation complex that contains histone acetylase activity. This provides a molecular mechanism for integrating chromatin remodelling and interactions with the core transcriptional machinery. □

Methods

Protein-protein interaction assays. Cells were lysed in NET-N buffer (20 mM Tris HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.5% NP40, 10% glycerol) containing protease inhibitors, clarified by centrifugation, and immunoprecipitated at 0°C. After incubation with protein-A/G agarose beads (Santa Cruz) at 4°C, the beads were washed with NET-N. Immunoprecipitated proteins were fractionated by SDS polyacrylamide gel electrophoresis, and either analysed by autoradiography or transferred to nitrocellulose for immunodetection using anti N-CoR (amino acids 1–312) guinea-pig IgG, horseradish peroxidase-conjugated anti-guinea-pig antibody (Sigma) and the ECL reagent system (Amersham). Anti-N-CoR (amino acids 2,239–2,453) was raised in rabbits and IgG fractions were prepared on a protein A FPLC column (Pharmacia). Interaction assays using GST fusion proteins were performed as described⁸.

Histone deacetylase assay. Assays for histone deacetylase activity were performed essentially as described^{47,48}. Immunoprecipitated complexes on protein A/G-sepharose beads were incubated for 30 min at 37°C with ~180 µg (150,000 d.p.m.) of acid-soluble histones isolated from ³H-acetate-labelled chicken erythrocytes. The reaction was terminated by adding acetic acid (to 0.12 M) and HCl (to 0.72 M). Released ³H-acetic acid was extracted with ethyl acetate and quantified by liquid scintillation counting. Samples were twice assayed in duplicate, and non-enzymatic release of label was subtracted to obtain the final value.

Nuclear microinjection, staining and fluorescence microscopy. Microinjection analysis was performed essentially as described^{47,48}. Before the injection, 293 cells or Rat-1 fibroblasts were rendered quiescent by incubation in serum-free medium for 24–36 h. Plasmids were injected into the nuclei of cells at a concentration of 100 µg ml⁻¹. Either preimmune rabbit IgG, or anti-N-CoR, anti-mRPD3 or various anti-mSin3 antibodies (Santa Cruz) were co-injected, allowing the unambiguous identification of the injected cells. Expression of proteins was allowed to proceed for at least 2 h before the addition of ligand, where indicated. β-Galactosidase activity was detected by incubation with X-gal. Injected cells were identified by staining with tetramethylrhodamine-conjugated donkey anti-rabbit IgG. In cells expressing high levels of β-galactosidase, the blue staining tended to quench rhodamine fluorescence. For this reason, injected cells were counted as those exhibiting either nuclear rhodamine fluorescence or blue X-gal staining or both. In many cases, cells injected with the GAL4 fusions of transcriptional repressors demonstrated a greatly decreased level of X-gal staining, a reflection of low-level expression of the reporter gene. Despite the obvious decrease in staining, all cells showing any trace of blue staining were scored as positive for expression, to avoid any possible subjectivity in the analysis.

Transfection assays. Transfections were performed as described⁸ using the reporters described in the figures. Essentially similar folds of repression were obtained in independent experiments in HeLa, Rat-1, Saos2 and CV-1 cell lines, but absolute numbers were higher for 293 cells and therefore data from these experiments are shown. Cells were treated with trapoxin for 24 h (Fig. 4d).

Received 22 January; accepted 31 March 1997.

1. Chambon, P. The retinoid signaling pathway: molecular and genetic analyses. *Semin. Cell Biol.* **5**, 115–125 (1994).
2. Mangelsdorf, D. J. *et al.* The nuclear receptor superfamily: the second decade. *Cell* **83**, 835–839 (1995).
3. Wong, J., Shi, Y. & Wolffe, A. P. A role for nucleosome assembly in both silencing and activation of the *Xenopus* TRBA gene by the thyroid hormone receptor. *Genes Dev.* **9**, 2696–2711 (1995).
4. Glass, C. K., Holloway, J. M., Devary, O. V. & Rosenfeld, M. G. The thyroid hormone receptor binds with opposite transcriptional effects to a common sequence motif in thyroid hormone and estrogen response elements. *Cell* **54**, 313–323 (1988).
5. Baniahmad, A., Kohne, A. C. & Renkawitz, R. A transferable silencing domain is present in the thyroid hormone receptor, in the v-erbA oncogene product and in the retinoic acid receptor. *EMBO J.* **11**, 1015–1023 (1992).
6. Baniahmad, A. *et al.* Interaction of human thyroid hormone receptor beta with transcription factor TFIIB may mediate target gene derepression and activation by thyroid hormone. *Proc. Natl Acad. Sci. USA* **90**, 8832–8836 (1993).
7. Baniahmad, A. *et al.* The T4 activation domain of the thyroid hormone receptor is required for release of a putative corepressor(s) necessary for transcriptional silencing. *Mol. Cell Biol.* **15**, 76–86 (1995).
8. Hörlein, A. J. *et al.* Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377**, 397–404 (1995).
9. Kurokawa, R. *et al.* Polarity-specific activities of retinoic acid receptors determined by a co-repressor. *Nature* **377**, 451–454 (1995).
10. Zamir, I. *et al.* A nuclear hormone receptor corepressor mediates transcriptional silencing by receptors with distinct repression domains. *Mol. Cell Biol.* **16**, 5458–5465 (1996).
11. Ayer, D. E., Kretzner, L. & Eisenman, R. N. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. *Cell* **72**, 211–222 (1993).
12. Ayer, D. E. & Eisenman, R. N. A switch from Myc:Max to Mad:Max heterocomplexes accompanies

- monocyte/macrophage differentiation. *Genes Dev.* **7**, 2110–2119 (1993).
13. Hurlin, P. J., Ayer, D. E., Grandori, C. & Eisenman, R. N. The Max transcription factor network. *Cold Spring Harb. Symp. Quant. Biol.* **59**, 109–116 (1994).
14. Ayer, D. E., Lawrence, Q. A. & Eisenman, R. N. Mac-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell* **80**, 767–776 (1995).
15. Schreiber-Agus, N. *et al.* An amino-terminal domain of Mx1 mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor SIN3. *Cell* **80**, 777–786 (1995).
16. Ayer, D. E., Laherty, C. D., Lawrence, Q. A., Armstrong, A. & Eisenman, R. N. Mad proteins contain a dominant transcriptional repression domain. *Mol. Cell Biol.* **16**, 5772–5781 (1996).
17. Nasmyth, K., Stillman, D. & Kipling, D. Both positive and negative regulators of HO transcription are required for mother-cell-specific mating type switching. *Cell* **48**, 579–587 (1987).
18. Sternber, P. W., Stern, M. J., Clark, I. & Herskowitz, I. Activation of the yeast HO gene by release from multiple negative controls. *Cell* **48**, 567–577 (1987).
19. Wang, H., Clark, I., Nicholson, P. R., Herskowitz, I. & Stillman, D. J. The *S. cerevisiae* SIN3 gene, a negative regulator of HO, contains four paired amphipathic helical motifs. *Mol. Cell Biol.* **10**, 5927–5936 (1990).
20. Vidal, M., Strich, R., Esposito, R. E. & Gaber, R. F. RPD1 is required for maximal activation and repression of diverse yeast genes. *Mol. Cell Biol.* **11**, 6306–6316 (1991).
21. Vidal, M. & Gaber, R. F. RPD3 encodes a second factor required to achieve maximal positive and negative regulation. *Mol. Cell Biol.* **11**, 6317–6327 (1991).
22. Wang, H. & Stillman, D. Transcriptional repression in *S. cerevisiae* by a SIN3-LexA fusion protein. *Mol. Cell Biol.* **13**, 1805–1814 (1993).
23. Nawaz, Z. *et al.* The yeast SIN3 gene product negatively regulates the activity of the human progesterone receptor. *Mol. Gen. Genet.* **245**, 724–733 (1994).
24. Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408–411 (1996).
25. Yoshida, M., Horinouchi, S. & Beppu, T. Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *BioEssays* **17**, 423–430 (1995).
26. Yang, W.-M., Inouye, C., Zeng, Y., Bearss, D. & Seto, E. Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc. Natl Acad. Sci. USA* **93**, 12845–12850 (1996).
27. Rose, D. W., McCabe, G., Feramisco, J. R. & Adler, M. Expression of c-fos and AP-1 activity in senescent human fibroblasts is not sufficient for DNA synthesis. *J. Cell Biol.* **119**, 1405–1411 (1992).
28. Sgouras, D. N. *et al.* ERF: an ETS-domain protein with strong transcriptional repressor activity, can suppress ets-associated tumorigenesis and is regulated by phosphorylation during cell cycle and mitogenic stimulation. *EMBO J.* **14**, 4781–4793 (1995).
29. O'Neill, E. M., Rebay, L., Tijian, R. & Rubin, G. M. The activities of two Ets-related transcription factors required for Drosophila eye development are modulated by the Ras/MAPK pathway. *Cell* **78**, 137–147 (1994).
30. Chen, J. D. & Evans, R. M. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature* **377**, 454–457 (1995).
31. Rundlett, S. E. *et al.* HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes. *Proc. Natl Acad. Sci. USA* **93**, 14503–14508 (1996).
32. Gray, S. & Levine, M. Transcriptional repression in development. *Curr. Opin. Cell Biol.* **8**, 358–364 (1996).
33. Kingston, R. E., Bunker, C. A. & Imbalzano, A. N. Repression and activation by multiprotein complexes that alter chromatin structure. *Genes Dev.* **10**, 905–920 (1996).
34. Svaren, J. & Horz, W. Regulation of gene expression by nucleosomes. *Curr. Opin. Cell Biol.* **6**, 164–170 (1994).
35. Hanna-Rose, W. & Hansen, U. Active repression mechanisms of eukaryotic transcription repressors. *Trends Genet.* **12**, 229–234 (1996).
36. Johnson, A. D. The price of repression. *Cell* **81**, 655–658 (1995).
37. Peterson, C. L. Multiple SWITCHes to turn on chromatin. *Curr. Opin. Cell Biol.* **6**, 171–175 (1996).
38. Roth, S. Y. & Allis, C. D. Histone acetylation and chromatin assembly: A single escort, multiple dances? *Cell* **87**, 5–8 (1996).
39. Wolffe, A. P. & Pruss, D. Targeting chromatin disruption. *Cell* **84**, 817–819 (1996).
40. Lee, D. Y., Hayes, J. J., Pruss, D. & Wolffe, A. P. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell* **72**, 73–84 (1993).
41. Seol, W., Mahon, M. J., Lee, Y. K. & Moore, D. D. Two receptor interacting domains in the nuclear hormone receptor corepressor RIP13/N-CoR. *Mol. Endocrinol.* **10**, 1646–1655 (1996).
42. Keleher, C. A., Redd, M., Schultz, J., Carlson, M. & Johnson, A. D. SSN6-Tup1 is a general repressor of transcription in yeast. *Cell* **68**, 709–719 (1992).
43. Almouzni, G., Khochbin, S., Dimitrov, S. & Wolffe, A. P. Histone acetylation influences both gene expression and development of *Xenopus laevis*. *Dev. Biol.* **165**, 654–659 (1994).
44. Ogryzko, V. V., Schlitz, R. L., Russanova, V., Howard, B. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
45. Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641–643 (1996).
46. Kamei, Y. *et al.* A CBP integrator complex mediates transcriptional activation and AP-1 inhibition by nuclear receptors. *Cell* **85**, 1–12 (1996).
47. Hendzel, M. J., Delcuve, G. P. & Davie, J. R. Histone deacetylase is a component of the internal nuclear matrix. *J. Biol. Chem.* **266**, 21936–21942 (1991).
48. Li, W., Chen, H. Y. & Davie, J. R. Properties of chicken erythrocyte histone deacetylase associated with the nuclear matrix. *Biochem. J.* **314**, 631–637 (1996).
49. Oñate, S. A., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science* **270**, 1354–1357 (1995).
50. Yang, X.-J., Ogryzko, V. V., Nishikawa, J.-I., Howard, B. H. & Nakatani, Y. *Nature* **382**, 319–324 (1996).
51. Laherty, C. D. *et al.* Cell (in the press).

Acknowledgements. T.H. and R.M.L. made equivalent contributions to this work. We thank A. Hörlein, A. Leon and J. Inostroza for plasmid reagents; M.-H. Ogilastro for N-CoR baculovirus construction; M. Yoshida for trapoxin; Santa Cruz Biotechnology for generating anti-mSin3 antibodies; C. Nelson for cell culture; S. Flynn for technical assistance; P. Meyer for in preparing illustrations; and B. Stawski for help in preparing this manuscript. T.H. is supported by the UC Breast Cancer Research Program; R.M.L. by fellowships from the Lucille Markey Foundation and the National Science Foundation; C.D.L. by the Ladies Auxiliary to the Veterans of Foreign Wars; J.T.T. by the National Cancer Institute of Canada; G.B. by a physician scientist training grant; and D.W.R. by an American Diabetes Association Career Development Award. R.N.E. is an American Cancer Society research professor. M.G.R. is an investigator with the Howard Hughes Medical Institute. These studies were supported by grants from the National Science Foundation to E.S. and the NIH to M.G.R., C.K.G. and R.N.E.

Correspondence and requests for materials should be addressed to M.G.R. (e-mail: mrosenfeld@ucsd.edu).

Role for N-CoR and histone deacetylase in Sin3-mediated transcriptional repression

Leila Alland^{*†}, Rebecca Muhle^{*}, Harry Hou Jr^{*}, Jason Potes^{*}, Lynda Chin^{*‡}, Nicole Schreiber-Agus^{*} & Ronald A. DePinho^{*}

^{*} Department of Microbiology and Immunology, [†] Department of Pediatrics, and [‡] Division of Dermatology, Albert Einstein College of Medicine, 1300 Morris Park Avenue C607, Bronx, New York, New York 10461, USA

Normal mammalian growth and development are highly dependent on the regulation of the expression and activity of the Myc family of transcription factors. Mxi1-mediated inhibition of Myc activities requires interaction with mammalian Sin3A or Sin3B proteins, which have been purported to act as scaffolds for additional co-repressor factors. The identification of two such Sin3-associated factors, the nuclear receptor co-repressor (N-CoR) and histone deacetylase (HD1), provides a basis for Mxi1/Sin3-induced transcriptional repression and tumour suppression.

Apart from mechanisms operating to ensure that there are appropriate levels of Myc mRNA and protein, the actions of Myc are controlled by its physical and functional interactions with other members of the Myc bHLH/LZ network; these include Myc's obligate DNA-binding partner Max and the Max-associated Mad/Mxi1 family (for review, see ref. 1). Mxi1 and Mad proteins act as potent antagonists of Myc through their competitive interaction with Max and occupation of Myc consensus DNA-binding sites, as well as their recruitment of mammalian homologues of a yeast transcriptional repressor Sin3 (also called RPD1 (refs 2, 3)). Loss of interaction with the mammalian proteins mSin3A and mSin3B, either by mutational disruption of the Mad/Mxi1 amino-terminal Sin3-interaction domain (SID)²⁻⁴ or by removal of the SID in a mouse Mxi1 isoform³, is associated with abolition of Mxi1 or Mad transcriptional repression and anti-oncogenic activity. The idea that the SID functions primarily to recruit an mSin3 molecule has been substantiated by the finding that mSin3A can substitute functionally for the Mxi1 SID in the inhibition of Myc activity⁵. Similarly, the transference of the SID onto heterologous DNA-binding proteins^{3,6,7} can convert them into potent suppressors of Myc/RAS cellular transformation^{3,6,7} and of Myc transactivation of reporters^{6,7}.

Like their yeast counterpart, mammalian Sin3 proteins possess four paired amphipathic helical (PAH) domains that are structurally analogous to the HLH motif, and are thought to mediate interactions between proteins⁸⁻¹⁰. Although little more is known about mSin3 proteins other than their nuclear localization⁵, clues as to their potential mechanism of action can be obtained from previous genetic studies of yeast Sin3 that demonstrated a functional relationship to SWI5. Genetic and biochemical studies indicate that members of the SWI family, in conjunction with those of the SNF family, disrupt chromatin structure during transcriptional activation (for review, see ref. 11). RPD3, identified with SIN3/RPD1 as a suppressor of the *swi5* mutation¹²⁻¹⁴, bears extensive homology to two mammalian histone deacetylases, human histone deacetylase HD1 (ref. 15) and mRPD3 (ref. 16), which are thought to be involved in nucleosomal condensation¹⁷. The finding that loss of either RPD3 or Sin3 leads to derepression of the same set of genes in *swi5* mutants¹³ suggests that they may operate in the same genetic pathway, although a physical interaction has yet to be documented. These and other yeast studies suggest that Sin3 interacts with essential DNA-binding proteins that place it in a promoter context in which it can repress transcription by recruiting proteins that condense chromatin architecture through deacetyl-

ation (for reviews see refs 17, 18). Beyond its possible involvement with chromatin regulators, the ability of Sin3 to function as a transcriptional repressor in reporter assays (in yeast cells)⁹ suggests that Sin3-associated proteins may also affect the activity of components of the basal transcriptional machinery. Although such models are appealing, the sole function of Sin3 as a repressor is inconsistent with the finding that the transcriptional activation of some genes is reduced in the *sin3* null mutants¹³.

We used yeast two-hybrid interaction screens and immunoprecipitation assays to identify candidate effectors of mSin3 transcriptional repression function. Our results led to the characterization of an interaction between mSin3 and the co-repressor N-CoR protein, a factor shown previously to mediate transcriptional silencing by nuclear hormone receptors^{19,20}. This interaction between mSin3 and N-CoR is shown to be necessary for mSin3-mediated repression of basal transcription. In addition, an association between mSin3 and human HD1 was mapped to the carboxy-terminal region of mSin3, a region that we show to be essential for Mxi1/mSin3-mediated suppression of cellular transformation. Multi-level regulation of Myc- (Mxi1-) responsive genes by the Mxi1/mSin3 complex on the level of alterations in chromatin structure may be accomplished in part through the differential recruitment of HD1 by alternative mSin3B forms.

N-CoR as an mSin3-associated protein

In the course of screening for the complete 3.2-kilobase mSin3B cDNA, which encodes 954 amino acids and all four PAH structures², we identified a ~2.1-kb cDNA that is predicted to encode a protein of 293 amino acids encompassing PAH1 and PAH2 only, designated mSin3B_{sf} (the short form of mSin3B). The first 274 amino acids are identical to those in the long-form mSin3B protein², designated here as mSin3B_{lf}, although the remaining 19 C-terminal amino acids in mSin3B_{sf} are unique. An analysis of the intron-exon organization at the junction where the short-form sequences diverge from those of the long form suggests alternative RNA processing of a single *msin3B* gene (data not shown). Expression studies demonstrate that both *msin3B* transcripts are broadly distributed and their proteins are found in the nucleus (data not shown). Alignment of amino-acid sequences of yeast and mammalian Sin3 proteins shows that the highest degree of similarity occurs in the PAH domains (in particular PAH1 and PAH2), as well as in the linker region between PAH3 and PAH4 (ref. 2). We suspected that some of these conserved regions were likely to act through the

recruitment of co-repressor proteins, and designed a combination of yeast two-hybrid and co-immunoprecipitation studies to verify this hypothesis.

A yeast two-hybrid screen of an embryonic mouse cDNA library²¹ with a mSin3A PAH1 bait yielded multiple His⁺/LacZ⁺ clones possessing cDNAs identical to N-CoR¹⁹. The portion of N-CoR in the rescued two-hybrid clones corresponds to two non-adjacent stretches in the published N-CoR sequence, suggesting that the two-hybrid isolate represents an alternatively spliced version of N-CoR (Fig. 1a). The N-CoR SID that engages the PAH1 domain of mSin3 (N-SID^{PAH1}) is distinct from the defined repression domains I and II and the nuclear receptor-interaction domain¹⁹.

Yeast small-scale transformation assays were used to confirm and localize more precisely the sites of interaction between N-CoR and mSin3. The segment of N-CoR isolated in the yeast two-hybrid screens (and the non-spliced form of N-CoR; not shown) interacted strongly and equally well with full-length mSin3B and with PAH1-containing fragments of mSin3B or mSin3A (Fig. 1b). No interaction with this specific region of N-CoR was observed with mSin3

segments that lacked the PAH1 domain (Fig. 1b) or with a variety of LexA fusion controls (not shown). Disruption of the presumed α -helices A and/or B of PAH1 by proline substitution was sufficient to abolish completely this specific interaction (Fig. 1b).

Bacterially expressed glutathione S-transferase (GST)-mSin3B fusion proteins were then tested for their ability to interact directly with various *in vitro* translated, radiolabelled N-CoR fragments. The *in vitro* translated N-CoR fragment bound to GST-mSin3B_{SF} but not to GST alone (Fig. 1c, lanes 1–3). When N-CoR fragments encompassing each side of the presumed splice junction (Fig. 1c, lanes 4–6 against lanes 7–9) were tested, only N-CoR (amino acids 1,681–1,893) was shown to interact with GST-mSin3B_{SF}, and deletion of a 13-residue α -helical structure within this N-CoR fragment eliminated this interaction (lanes 10–12). This specific N-CoR region did not bind to GST-mSin3B (PAH3), GST-mSin3B (PAH4), GST alone or GST fusions to unrelated proteins (lanes 16–20). Reciprocally, a substitution of proline for leucine in helix A of mSin3B PAH1 adversely affected the ability of N-CoR to bind GST-mSin3B_{SF} (compare in Fig. 1c, lane 15 to lane 14). Together, these

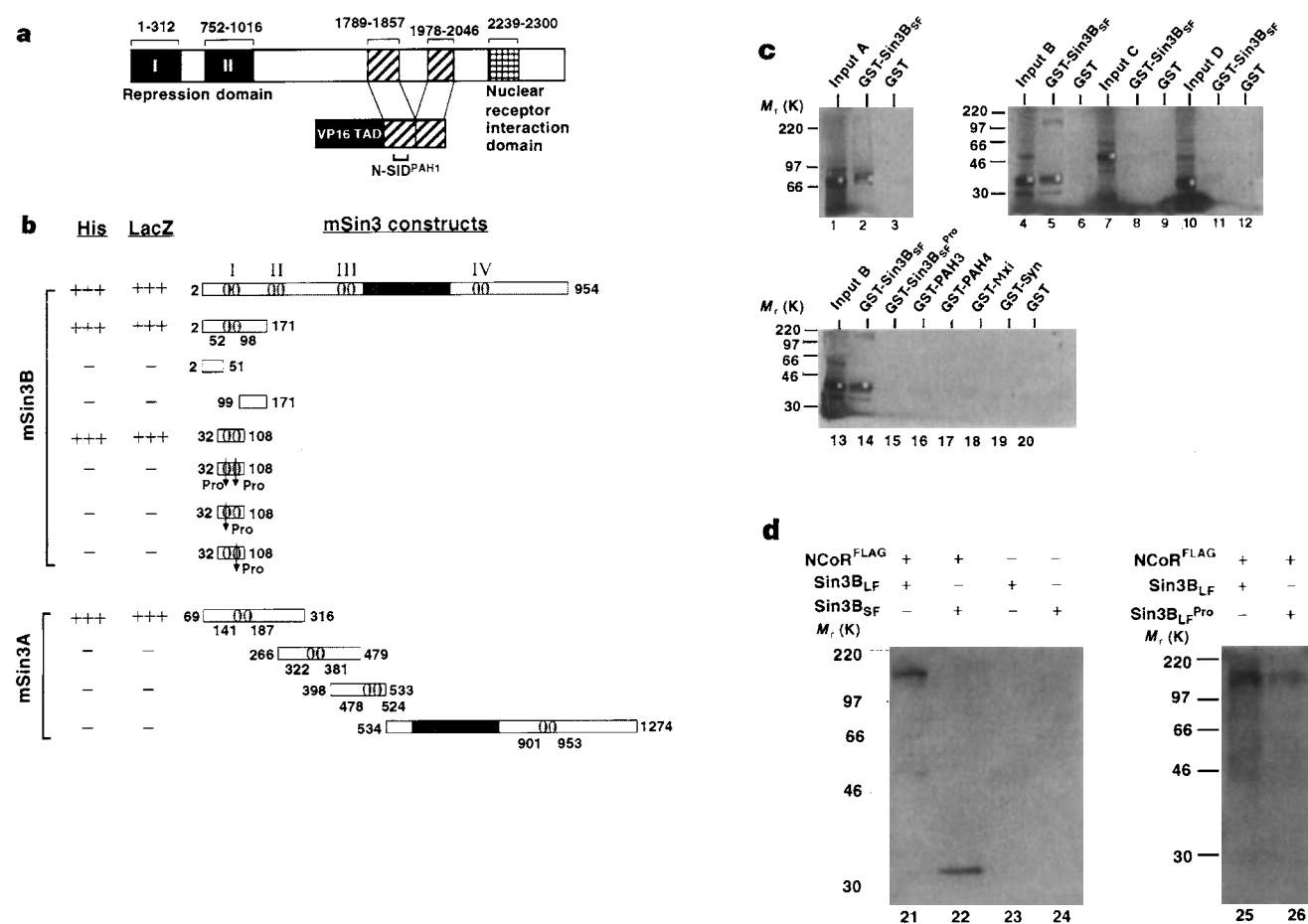


Figure 1 Localization of sequences required for interaction between N-CoR and mSin3A or mSin3B. **a**, Structure-function features of the previously reported N-CoR and its topological relationship to the VP16(TAD)-N-CoR fusion identified in our yeast two-hybrid screens. **b**, Various mSin3B and mSin3A baits used along with the VP16(TAD)-N-CoR activator to localize the N-CoR interaction domain of mSin3 in yeast cells. + + +, a His⁺ and LacZ⁺ phenotype that is equivalent to that observed for Myc-Max HLH/LZ interaction in yeast; -, no growth in the absence of His and no LacZ activity. Protein expression of the various baits was verified by western blot (data not shown). **c**, Localization of N-SID^{PAH1} by GST pull-down assays. The GST fusions are listed above each lane. Each lane has been programmed with a radiolabelled transcribed/translated product: lanes 1–3, N-

CoR 1,681–2,218 'input A'; lanes 4–6, N-CoR 1,681–1,893 'input B'; lanes 7–9, N-CoR 1,892–2,218 'input C'; lanes 10–12, N-CoR 1,681–1,893 bearing a deletion of residues 1,833–1,845 'input D'; and lanes 13–20, N-CoR 1,681–1,893 'input B'. Input, 10%. **d**, *In vivo* expression of N-CoR^{FLAG} along with various mSin3B expression constructs immunoprecipitated with anti-flag M2 antibody results in co-immunoprecipitation of mSin3B_{LF} (lane 21) or mSin3B_{sf} (lane 22), detected after western blotting using anti-mSin3B PAH2 antibody. Co-immunoprecipitation of mSin3B_{LF} bearing proline mutations in helices A and B of PAH1 was diminished compared to mSin3B_{LF} (lanes 25 and 26). Expression levels of mSin3B_{LF} and mSin3B_{LF}^{Pro} were similar (data not shown). The *M_r* values of immunoprecipitated proteins were 114K for mSin3B_{LF} and 33K for mSin3B_{sf}.

findings indicate that putative α -helical structures in each protein (13 residue N-SID^{PAH1} and PAH1 helices) are required for this interaction. Low-stringency co-immunoprecipitation assays following the transient transfection of mammalian cells verified that N-CoR can interact *in vivo* with both long and short forms of mSin3B. Additionally, double proline substitutions in PAH1 of mSin3B_{LF} diminished but did not eliminate N-CoR association (Fig. 1d), indicating that there are at least two distinct points of contact between mSin3B and N-CoR *in vivo*.

Transcriptional repression by mSin3B

Previous studies have demonstrated that yeast Sin3 and the mammalian Mad/mSin3 complex exhibit transcriptional repression activity in reporter assays^{2,9}. Our reporter assays show that both mSin3B_{LF} and mSin3B_{SF}, when fused to the GAL4 DNA-binding domain (GAL4), can repress basal transcription of a chloramphenicol acetyl transferase (CAT) reporter bearing five GAL4 DNA-binding sites upstream of either the major late promoter (MLP) (Fig. 2a; lane 2, mSin3B_{LF}; lane 4, mSin3B_{SF}) or thymidine kinase (TK) promoter (Fig. 2b, lanes 7 and 12). Similar results were obtained when full-length mSin3A, or the region of mSin3A containing PAH1 and PAH2 only, was fused to GAL4 (data not shown). Analogous to the requirement of nuclear hormone receptor association with N-CoR in reporter gene repression¹⁹, the disruption of helix A in PAH1 results in a substantial attenuation in the repression by the GAL4-mSin3B fusions (Fig. 2a; mSin3B_{LF}, compare lane 3 to 2; mSin3B_{SF}, compare lane 5 to 4). However, derepression is not complete, as indicated by the residual repression activity of the proline mutants (Fig. 2a; compare lanes 3 and 5 to 1) and by the partial repression of GAL4-Sin3B(PAH3 + 4) (Fig. 2a; compare lane 6 to 1), findings in agreement with the protein interaction data above indicating additional points of contact between N-CoR and mSin3B beyond the PAH1-N-SID^{PAH1} association. These results are analogous to reporter studies in yeast demonstrating that complete loss of Sin3 repression requires deletion of PAH3 and at least one other PAH domain⁹.

The physiological importance of the mSin3 PAH1 interaction with N-CoR in Sin3-mediated repression was verified by the capacity of a VP16 transactivation domain-N-CoR SID fusion protein (VP16-N-SID^{PAH1}) to alleviate PAH1-dependent repression. Addition of VP16-N-SID^{PAH1} to GAL4-mSin3B_{LF} or GAL4-

mSin3B_{SF} transfections led to increased CAT activity (Fig. 2b, compare lane 8 to 7, and lane 13 to 12). Furthermore, consistent with the existence of multiple N-CoR-mSin3B contact points, this derepression was more muted in mSin3B_{LF} than in mSin3B_{SF} (Fig. 2b, compare lane 8 to 13). In contrast, VP16-NSID^{PAH1} had no effect on the activity of mSin3B_{LF}^{Pro} (lanes 10 and 11). Finally, derepression by VP16-N-SID^{PAH1} was abolished with deletion of the α -helix in the N-SID^{PAH1} portion of the fusion protein (VP16- Δ NSID, lanes 9 and 14).

Repression of Myc transformation by mSin3B

We have shown previously that the addition of Mxi1 bearing its SID (Mxi-SR) to Myc/RAS co-transfections dramatically reduces the number of transformed foci generated in the highly quantitative REF cooperation assay³. The development of an analogous structure-function assay for mSin3 is complicated by: high endogenous mSin3 levels in REFs that mask the effects of exogenously introduced mSin3 (ref. 5); the fact that mSin3 proteins act on Myc-responsive gene targets *in trans*^{2,3}; and the possibility that mSin3 proteins interact with other sequence-specific DNA-binding proteins. These considerations were taken into account in the design of a biological assay in which the full-length open reading frame (ORF) of mSin3B_{LF} or engineered derivatives is fused in-frame to the Mxi1 ORF just downstream of the SID (Fig. 3a) and tested for the ability to suppress Myc/RAS cotransformation.

Independent experiments demonstrated that addition of the mSin3B_{LF}-Mxi expression construct to Myc/RAS co-transfections led to a consistent 2.6- to 4-fold reduction in the number of transformed foci (Fig. 3b; compare vector only to mSin3B_{LF}-Mxi). This level of suppression was comparable to that observed for Mxi-SR (Fig. 3b; Mxi-SR, 4- to 5-fold)³ and that reported previously for a similarly designed full-length mSin3A-Mxi fusion (4- to 5-fold)³. As well as reduced foci numbers, the anti-oncogenic effect of mSin3B_{LF}-Mxi was evident from a delay in the appearance of Myc/RAS/Sin3B_{LF}-Mxi foci in the monolayer (increased latency), less transformed morphology of these foci (data not shown), and significant reduction in their subcloning efficiency in comparison to Myc/RAS controls (Fig. 3c; 89% for the Myc/RAS control compared with 46% for mSin3B_{LF}-Mxi). In sharp contrast, addition of the mSin3B_{SF}-Mxi fusion construct to Myc/Ras co-transfections had a

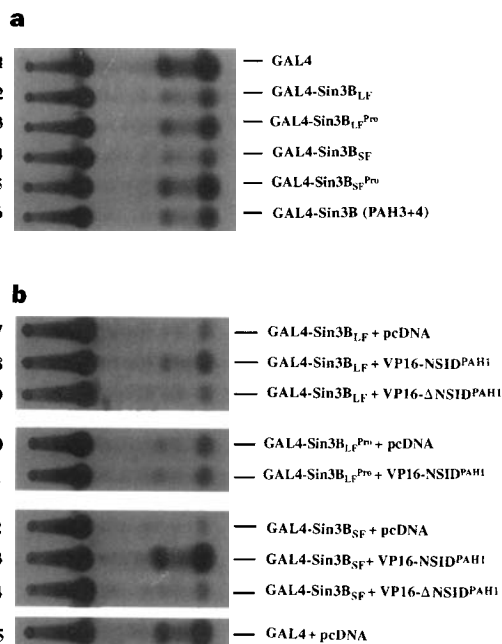


Figure 2 GAL4-mSin3B-mediated repression of basal transcription of CAT reporters bearing five GAL4 DNA-binding sites upstream of the major late promoter or thymidine kinase promoter. **a**, Various GAL4-mSin3B fusions were tested for their ability to repress basal transcription. Average levels of repression compared to GAL4 vector (lane 1) were GAL4-mSin3B_{LF} (lane 2), 11-fold; GAL4-mSin3B_{LF}^{Pro} (lane 3), 5-fold; GAL4-mSin3B_{SF} (lane 4), 14-fold; GAL4-mSin3B_{SF}^{Pro} (lane 5), 2-fold; GAL4-mSin3B (PAH3 + 4) (lane 6), 5-fold. Western blot analysis of cell lysates following transient transfection of the various GAL4-mSin3B fusions indicated similar expression levels (data not shown). **b**, Coexpression of a VP16 transactivation domain (VP16)-NSID^{PAH1} fusion protein alleviates GAL4-mSin3B-mediated repression. GAL4-mSin3B_{LF} (lanes 7-9), GAL4-mSin3B_{LF}^{Pro} (lanes 10, 11) and GAL4-mSin3B_{SF} (lanes 12-14) were co-transfected along with empty vector (lanes 7, 10 and 12), VP16-NSID encoding the VP16 transactivation domain fused to N-CoR residues 1,681-2,218 (lanes 8, 11 and 13), or VP16 Δ NSID encoding the VP16 transactivation domain fused to N-CoR residues 1,681-2,218 bearing a deletion of residues 1,833-1,845 (lanes 9 and 14), and compared with GAL4 alone (lane 15) for their ability to repress basal transcription.

minimal effect on foci number, latency, morphology or subcloning efficiency (Fig. 3b, Sin3B_{SE}-Mxi; and Fig. 3c, 89% for Myc/RAS versus 88% for mSin3B_{SE}-Mxi), similar to that reported previously for Mxi-WR (a naturally occurring alternative Mxi1 form that lacks the SID) or for Mxi-SR^{Pro} (an engineered mutant that is unable to interact with mSin3)³.

The apparent lack of mSin3B_{SE} anti-oncogenic activity indicated that sequences present only in mSin3B_{LF} are essential and possibly sufficient for anti-oncogenic activity. To examine whether the portion of mSin3B_{LF} that corresponds to mSin3B_{SE} contributes at all to the anti-oncogenic activity of mSin3B_{LF}, we tested the activity of mSin3B_{LF}-Mxi fusions containing either proline substitutions (in both PAH1 helices) or an in-frame deletion of the PAH1 + 2 region. The proline substitution did not significantly affect the anti-oncogenic activity of the mSin3B_{LF}-Mxi fusion (Fig. 3b; mSin3B_{LF}^{Pro}-Mxi; 4-fold reduction in comparison to empty vector). Although the PAH1 proline mutant of mSin3B_{LF} can still associate with N-CoR *in vivo* (Fig. 1d), these results suggest that the N-CoR/PAH1 interaction does not play a significant role in suppression of cellular transformation, and additional co-factors that associate with the region C-terminal to PAH2 (present in mSin3B_{LF} and mSin3A but not mSin3B_{SE}) are needed to achieve adequate tumour-suppression activity. This view is supported further by the ability of a mSin3B_{LF} mutant, deleted for PAH1 and PAH2 regions and designated mSin3B(PAH3 + 4)-Mxi, to reduce Myc/RAS foci formation by an average of 2.5- to 3.5-fold (Fig. 3b; mSin3B(PAH3 + 4)-Mxi). The slightly less efficient repression effected by this deletion construct in comparison to mSin3B_{LF} could relate to structural anomalies induced by the truncation, loss of synergy with co-factors interacting with the PAH1 + 2 region, or diminished N-CoR interaction, among other possibilities, and does not appear to result from variances in protein stability (data not shown).

HD1 interacts with mSin3B_{LF} but not mSin3B_{SE}

The above results suggest that the region C-terminal to PAH2 is integral to mSin3B-mediated tumour suppression. The link between histone acetylation factors and cancer-prone conditions, together with genetic data from yeast implicating a tight functional relationship between Sin3 and RPD3 (the yeast homologue of the human histone deacetylase, HD1)¹², prompted us to assess possible physical interactions between HD1 and this functionally important region of mammalian Sin3B_{LF}. Accordingly, low-stringency co-immunoprecipitation was used to determine whether HD1 and mSin3B interact physically in mammalian cells and, more specifically, to establish whether HD1 interacts with the long form of mSin3B (PAH1 + 4) rather than the short form (PAH1 + 2). mSin3B immunoprecipitates of cells singly transfected with mSin3B_{LF} identified two bands corresponding to the predicted sizes of endogenous HD1 (M_r 55K) and mRPD3 (M_r 59K) (lane 4). In co-transfections, anti-FLAG antibody immunoprecipitated mSin3B_{LF}, but not mSin3B_{SE}, when each mSin3B construct was co-transfected separately with HD1^{FLAG} (Fig. 4a, lanes 8 and 10, respectively). Next, we assessed whether the functionally active mSin3B(PAH3 + 4)-Mxi fusion protein was also capable of association with HD1 (Fig. 4b). Anti-FLAG immunoprecipitations yielded the HD1^{FLAG} protein (lanes 11, 13 and 14) along with its protein partner of expected size: mSin3B(PAH3 + 4)-Mxi (lane 13) or mSin3B_{LF} control (lane 14). These results demonstrate that the PAH3 + 4 region, which possesses potent anti-oncogenic activity, participates in HD1 association, and is sufficient for the interaction between HD1 and mSin3B when linked directly to Mxi1. Finally, low-stringency immunoprecipitation with either anti-FLAG or anti-Myc failed to detect association between HD1^{FLAG} and Mxi-SR^{Myc} in cells doubly transfected with HD1^{FLAG} and Mxi-SR^{Myc} expression constructs (Fig. 4c, lanes 17 and 18). In

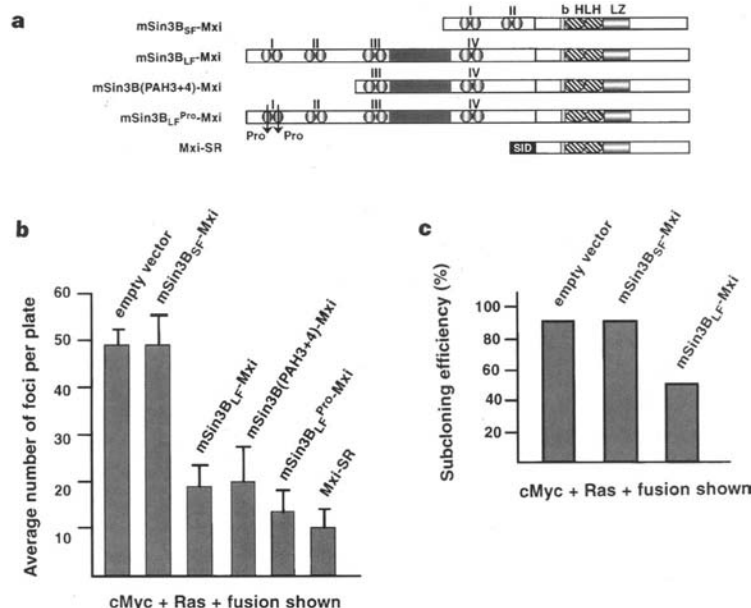


Figure 3 The anti-Myc activity of Mxi-SR in the REF co-transformation assay is mediated by sequences in the C-terminal PAH3 and PAH4 region of mSin3B_{LF}. **a**, Schematic diagram of the inserts that were subcloned into the pVNic vector³ to generate the expression constructs used in the REF assay. The mSin3B_{SE}-Mxi, mSin3B_{LE}-Mxi and mSin3B(PAH3 + 4)-Mxi fusions were generated by ligating the appropriate mSin3B sequences in-frame with a PCR-generated fragment of Mxi-SR that was deleted for most of the N-terminal SID domain, represented here by a black box. mSin3B_{LE}^{Pro} contained substitutions of proline for leucine in PAH1 helices A and B. **b**, Early passage cultures of REFs were co-transfected with myc,

ras and an expression construct listed above the bars and depicted in **a** as described³. The total number of foci on six plates was counted approximately 10 days after transfection. Average number of foci per plate from a representative experiment is shown; error bars indicate s.d. **c**, Foci from transfected points myc/ras/empty vector, myc/ras/mSin3B_{SE}-mxi and myc/ras/mSin3B_{LE}-mxi were picked individually and expanded for the establishment of transformed cell lines. Subcloning efficiency is expressed as the percentage of foci that gave rise to established lines.

contrast, the appearance of Mxi-SR^{Myc} in anti-FLAG immunoprecipitates of cells triply transfected with Mxi-SR^{Myc}, mSin3B_{LF} and HD1^{FLAG} (lane 19) indicates that the three proteins are components of the same multi-protein complex *in vivo*.

Discussion

Mechanistic insight into how mSin3 participates as an obligate partner in an Mxi1 repression pathway has now been obtained through the identification of two mSin3-associated proteins that are established modulators of gene expression: N-CoR and HD1. These findings support the idea that multiple co-repressors may jointly bring about repression mediated by a single DNA-binding transcription factor. In addition, the existence of two mSin3B forms suggests that distinct types of Mxi1-mediated repression could occur through the recruitment of different combinations of these co-repressor molecules. The recruitment of N-CoR (but not HD1) by mSin3B_{SF} could provide a more attenuated and readily reversible

type of regulation operating primarily on the basal transcription apparatus. In contrast, the ability of mSin3B_{LF} to associate with N-CoR as well as HD1, the latter a histone deacetylase and presumed nucleosomal condenser¹⁷, may allow for a more effective multi-level means of gene repression. This situation is reminiscent of the yeast global repressors Ssn6 and Tup1 which use multiple mechanisms to achieve full repression *in vivo*²²⁻²⁴. Although Ssn6 and Tup1 are able to repress transcription *in vitro* in reporter assays (presumably in the absence of chromatin)²⁵, the extent of repression (only 4-fold) is much lower than that seen *in vivo* (over 100-fold), suggesting a level of regulation that extends beyond inhibition of the basal transcription machinery and could involve the modulation of higher-order structures of promoter DNA. Similarly, chromatin assembly has been documented to play an important role in the ability of thyroid receptor to repress basal transcription of the *Xenopus TRβA* gene²⁶.

The existence of two forms of mSin3B, both with the capacity to interact with the Mxi1(Mad) SID, raises the question of whether

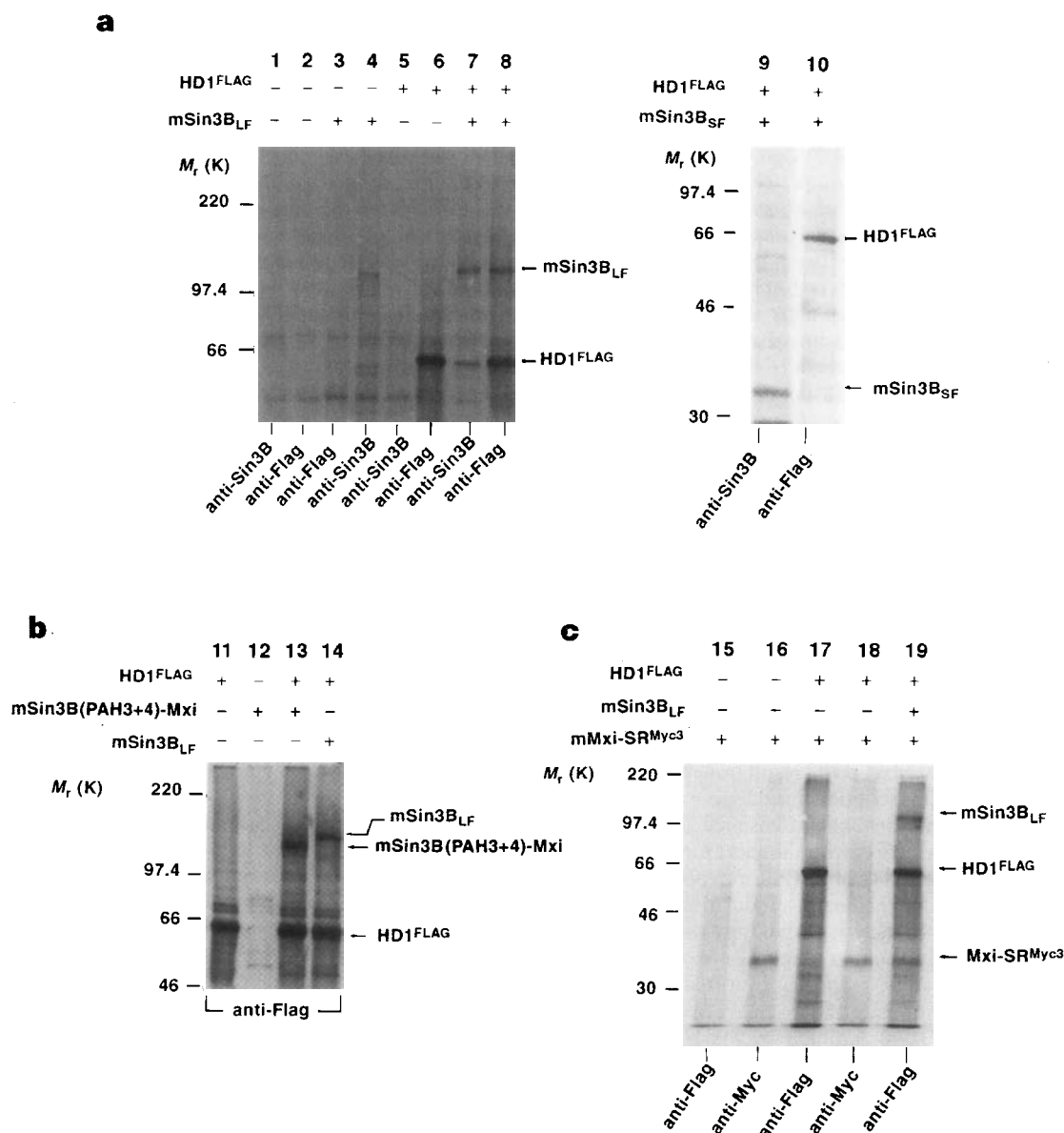


Figure 4 mSin3B_{LF} interacts with HD1 *in vivo*. 293T cells were metabolically labelled after co-transfection of HD1^{FLAG} and the various mSin3B expression constructs shown. Radiolabelled whole-cell extracts were immunoprecipitated under low-stringency conditions using either anti-mSin3B antibody, anti-FLAG

antibody M2, or anti-Myc antibody, and visualized by autoradiography. The relative *M_r* values of immunoprecipitated proteins were 59K for HD1^{FLAG}, 114K for mSin3B_{LF}, 33K for mSin3B_{SF}, 103K for mSin3B(PAH3 + 4)-Mxi, and 35K for Myc-tagged Mxi.

mSin3B_{SE} could function to attenuate chromatin regulation by Mxi1(Mad) by blocking mSin3B_{LF} and HD1 recruitment to promoters of growth-control genes. Attenuation by Sin3B_{SE} could occur in times of active growth and differentiation, when maximal chromatin relaxation might be necessary for effective yet controlled transcription of Myc-responsive developmental genes. In contrast, in settings such as terminal differentiation or senescence, there may be a need to downregulate target genes involved in active growth and development through multiple and functionally distinct repression mechanisms. Such mechanisms would include not only the downregulation of *c-myc* gene expression and the concomitant upregulation of *mxil* (as observed in differentiating erythrocytes)³, but also the loss of the putative role of mSin3B_{SE} as an attenuator, thereby allowing mSin3B_{LF} to regulate chromatin. A similar situation of differential splicing yielding protein products with distinct effects on cell growth has been described for the mammalian Brg proteins, which are homologues of the yeast Swi2/Snf2 complexes^{27–29}. The distinct effects of the Brg proteins may relate to their different abilities to modulate chromatin structure^{30,31}.

In this study, our use of a gene-fusion assay to map regions of functional importance in mSin3B has implicated the HD1 interaction domain of mSin3B_{LF} in Mxi1-mediated tumour suppression. Although it is possible and likely that additional mSin3B-associated proteins are required for suppression of the transformation process, many observations have supported a significant role for chromatin-modulating factors in the genesis or suppression of cancer. Some notable examples include the leukaemia-associated chromosomal translocation event fusing histone acetyltransferase CBP and MOZ^{32,33}, and the strong correlation between E1a oncogenic competence and its ability to disrupt the growth-inhibitory interaction between the histone acetyltransferases p300 and P/CAF³⁴. Most relevant to the Myc pathway is the finding that the *bmi-1* proto-oncogene can collaborate with Myc to induce B-cell lymphoma in *Em-myc* transgenic mice^{35,36}. The Bmi-1 protein shares homology with the *Drosophila* protein Posterior Sex Combs, a member of the Polycomb group involved in maintaining a stable repressed chromatin structure over the promoters of homoeotic genes during development³⁷. The emerging role for chromatin modulators in oncogenic mechanisms lends support to the view that the anti-oncogenic activity of the mSin3B(PAH3 + 4) domain is linked to its ability to recruit histone deacetylase activity.

The construction of a simple and generalized model of the functional relationship between Mxi1 and Myc at the level of commonly regulated transformation genes is complicated by the fact that Myc can operate both as repressor (for example, albumin, C/EBP) and activator (for example, ODC, Cdc25A) of gene expression^{1,38–42}. If Mxi1 functions to repress Myc-activated targets, the anti-oncogenic actions of Mxi1 could result from N-CoR/HD1-mediated inhibition of the transcriptional machinery and deacetylation of core histones with resultant nucleosomal condensation. The lack of an anti-oncogenic effect with the short form of mSin3B could be explained by the inability of N-CoR to repress effectively such targets in the absence of chromatin-modulating activities of HD1 (somewhat analogous to the yeast Ssn6/Tup1 repressor data). Alternatively, Mxi1 may also act to reverse Myc repression of growth-arrest genes. This can be supported by a less conventional view that histone deacetylation may lead to a more open chromatin structure, thus making HD1 recruitment by Mxi1/mSin3B_{LF} an activating event. This view is in contrast to that generally held, that histone deacetylation correlates with repression of gene expression^{43,44} and that histone acetylation enhances *trans*-acting factor access to nucleosomal DNA⁴⁵. However, the anti-oncogenic effects of histone deacetylase inhibitors and genetic studies in yeast and *Drosophila* support the view that HD1 can have an activating function^{46,47}. Specifically, the recent molecular/genetic analysis of a null *rpds* mutant has revealed an unexpected enhancement of position–effect variegation in *Drosophila* and telomeric position–

effect in yeast⁴⁷, suggesting that histone deacetylase is involved in counteracting genomic silencing by derepression of such loci. Thus the HD1 recruitment by Mxi1/mSin3B_{LF} to Myc-responsive targets, together with a downregulation of the suppressive functions of N-CoR on the basal machinery, could act to derepress Myc target genes involved in growth suppression pathways. Notwithstanding this possible ‘activating’ role for HD1 in the Myc transformation pathway, our data do not exclude the possibility that, in some promoter contexts or given a particular combination of histone H4 isoforms, HD1 in conjunction with N-CoR may serve a role as ‘repressor’ that counteracts Myc’s activation of growth-promoting genes. The targeted delivery of N-CoR and/or HD1 in the form of Mxi fusions reported here may provide a system for the study of activating versus repressing functions in a promoter-specific manner.

Finally, the interaction of N-CoR with components of both the nuclear hormone receptor family and the Myc superfamily raises the possibility that these pathways may be functionally integrated. Integration could be based upon competition between nuclear hormone receptors (such as the retinoic-acid receptor) and Mxi1(Mad) for limiting amounts of the co-repressors and resultant coordinate regulation of genes in the nuclear hormone receptor and the Myc superfamily pathways. Along these lines, it is noteworthy that many retinoic acid-responsive cell-culture differentiation systems (for example, HL-60 and Neuroblastomas) are also highly Myc dependent⁴⁸. In these model systems, exposure of cells to retinoic acid could result in activation of retinoic-acid receptor-responsive genes through the recruitment of co-activators and the coordinate release of N-CoR. The increased pool of N-CoR would then be available for interaction with the differentiation-associated rising levels of Mxi1, and allow the repression of genes necessary for the progression towards a terminally differentiated state. □

Methods

Isolation of mSin3B_{SE} cDNA. A cDNA encoding mSin3B_{SE} was isolated by screening an amplified newborn mouse brain cDNA library (Stratagene) with a previously reported two-hybrid-derived fragment encoding the PAH2 domain of mSin3B^{2,3}. The full-length mSin3B cDNA was provided by R. N. Eisenman.

Yeast two-hybrid screens. A mouse embryonic day 9.5 and 10.5 cDNA library in the vector pVP16 (provided by S. Hollenberg)²¹ was introduced into the *Saccharomyces cerevisiae* L40 reporter strain expressing LexA fused to the PAH1 domain of mSin3A (LexA–mSin3A(PAH1)). The LexA plasmid, pBTM116, was provided by R. Sternglanz. Plasmids encoding the VP16 TAD fused to various regions of N-CoR were introduced along with plasmids encoding LexA fused to various regions of mSin3A or mSin3B into the L40 yeast strain by small-scale transformation. Substitutions of proline for leucine in helices A (L59P) and/or B (L89P) of PAH1 for this and other experiments were introduced by PCR-based site-directed mutagenesis. Double transformants were assayed for histidine prototrophy and for β-galactosidase activity by filter assay²¹.

In vitro translation and GST pull-down assays. The two-hybrid-derived N-CoR cDNA insert was used to probe a newborn mouse brain cDNA phage library (Stratagene). Restriction mapping and sequence analysis of several N-CoR-hybridizing clones resulted in the identification of a partial cDNA clone, pN-CoR1, containing a portion of the published N-CoR ORF. Various pN-CoR1 fragments were cloned into the coupled transcription/translation vector pRSET (Invitrogen). PCR-based site-directed mutagenesis was used to delete the predicted α-helical structure (Δ1,833–1,845). The GST fusion constructs were generated by fusing fragments of mSin3B, full-length Mxi-SR³, or syndecan-2 (H. Shen and R.A.D., in preparation) cDNAs in-frame into pGEX (Pharmacia). GST–mSin3B_{SE}, GST–mSin3B(PAH3), and GST–mSin3B(PAH4) encode residues 2–274, 298–390 and 687–888 of mSin3B, respectively. GST fusion proteins were expressed in DH5-α cells and purified with glutathione-Sepharose beads (Pharmacia). Radiolabelled TNT-coupled transcription/translation products (Promega) were incubated with the various GST fusion proteins for 1 h at 4°C in phosphate-buffered saline containing an additional 100 mM KCl and 0.5% NP-40 (L100) (ref. 49). Glutathione-Sepharose beads were added to the binding reaction for an

additional 30 min, then washed five times in L100. Bound proteins were eluted with sample buffer, fractionated on SDS polyacrylamide gels, and visualized by autoradiography.

Immunoprecipitations. Subconfluent 293T cells were transfected with 3–5 µg each of the appropriate expression constructs shown in Figs 1d and 4, and 80 µg of Lipofectamine reagent (Gibco BRL). Immunoprecipitations under low-stringency conditions were performed³ using anti-FLAG M2 (Kodak), anti-mSin3B (PAH2) AK12 (Santa Cruz), and anti-Myc OP10 (Calbiochem) antibodies. To study HD1 interactions, cells were metabolically labelled using the EXPRE^{35S} protein-labelling mix (Dupont-NEN) for 6 h before collection. To generate the mSin3B_{LF} and mSin3B_{SF} expression constructs, mSin3B cDNAs encoding residues 1–954 and 1–274, respectively, were ligated into pcDNA (Invitrogen). The full-length N-CoR^{FLAG} expression construct was provided by M. G. Rosenfeld. The HD1^{FLAG} expression construct in pBJ5 was provided by S. Schreiber. The myc-tagged Mxi-SR construct in pJFE14 was described previously³.

Reporter assays and western blot. Subconfluent NIH3T3 cells were co-transfected with 3 µg of various GAL4–mSin3B fusion constructs and 14 µg of the G5MLPCAT or 18 µg of the G5TKCAT reporter³⁰. Transfection efficiencies were determined by addition of 2 µg of human growth hormone plasmid to each transfection point and assaying media for human growth hormone by radioimmunoassay (Nichols Institute) just before cell lysis. Normalized amounts of cell extracts were analysed for CAT activity by standard methods. The GAL4 (DBD only) construct in the pGALO vector was provided by C. Dang. G5MLPCAT and G5TKCAT reporters were provided by A. J. Berk. The GAL4–mSin3 constructs were cloned into pGALO and encode the GAL4 DBD fused to various mSin3B segments including mSin3B_{LF} (residues 2–954), mSin3B_{SF} (residues 2–274), and the region spanning PAH3 and PAH4 of mSin3B_{LF} (residues 298–954).

REF cooperation assays. Early-passage rat embryo fibroblast monolayers were prepared and co-transfected by the calcium phosphate precipitation method³. To make the various mSin3B–Mxi fusion expression constructs, mSin3B_{LF} (encoding residues 1–954) or mSin3B_{SF} (encoding residues 1–274) was ligated in-frame to Mxi-SR deleted for the SID (MxiΔrep)³ producing mSin3B_{LF}–Mxi and mSin3B_{SF}–Mxi, respectively. The mSin3B(PAH3 + 4)–Mxi expression construct was generated by replacing a fragment of the mSin3B_{LF}–Mxi construct with bridging oligonucleotides such that mSin3B_{LF} residues 12–293 were deleted.

Received 23 January; accepted 31 March 1997.

- Henriksson, M. & Luscher, B. Proteins of the Myc network: essential regulators of cell growth and differentiation. *Adv. Cancer Res.* **68**, 109–182 (1996).
- Ayer, D. E., Lawrence, Q. A. & Eisenman, R. N. Mad-Max transcriptional repression is mediated by ternary complex formation with mammalian homologs of yeast repressor Sin3. *Cell* **80**, 767–776 (1995).
- Schreiber-Agus, N. *et al.* An amino-terminal domain of Mxi1 mediates anti-Myc oncogenic activity and interacts with a homolog of the yeast transcriptional repressor Sin3. *Cell* **80**, 777–786 (1995).
- Koskinen, P. J., Ayer, D. E. & Eisenman, R. N. Repression of Myc-Ras cotransformation by Mad is mediated by multiple protein-protein interactions. *Cell Growth Differ.* **6**, 623–629 (1995).
- Rao, G. *et al.* Mouse Sin3A interacts with and can functionally substitute for the amino-terminal repression domain of the Myc antagonist Mxi. *Oncogene* **12**, 1165–1172 (1996).
- Harper, S. E., Qiu, Y. & Sharp, P. A. Sin3 corepressor function in Myc-induced transcription and transformation. *Proc. Natl Acad. Sci. USA* **93**, 8536–8540 (1996).
- Ayer, D. E., Laherty, C. D., Lawrence, Q. A., Armstrong, A. P. & Eisenman, R. N. Mad proteins contain a dominant transcription repression domain. *Mol. Cell. Biol.* **16**, 5772–5781 (1996).
- Wang, H., Clark, I., Nicholson, P. R., Herskowitz, I. & Stillman, D. J. The *Saccharomyces cerevisiae* SIN3 gene, a negative regulator of HO, contains four paired amphipathic helix motifs. *Mol. Cell. Biol.* **10**, 5927–5936 (1990).
- Wang, H. & Stillman, D. J. Transcriptional repression in *Saccharomyces cerevisiae* by a SIN3-LexA fusion protein. *Mol. Cell. Biol.* **13**, 1805–1814 (1993).
- Vidal, M., Strich, R., Esposito, R. E. & Gaber, R. F. RPD1 (SIN3/UME4) is required for maximal activation and repression of diverse yeast genes. *Mol. Cell. Biol.* **11**, 6306–6316 (1991).
- Winston, F. & Carlson, M. Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. *Trends Genet.* **8**, 387–391 (1992).
- Nasmyth, K., Stillman, D. J. & Kipling, D. Both positive and negative regulators of HO transcription are required for mother-cell-specific mating-type switching in yeast. *Cell* **48**, 579–587 (1987).
- Vidal, M. & Gaber, R. F. RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **11**, 6317–6327 (1991).
- Stillman, D. J., Dorland, S. & Yu, Y. Epistasis analysis of suppressor mutations that allow HO expression in the absence of the yeast SWI5 transcriptional activator. *Genetics* **136**, 781–788 (1994).
- Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast

transcriptional regulator Rpd3p. *Science* **272**, 408–411 (1996).

- Yang, A.-M., Inouye, C., Zeng, Y., Bearss, D. & Seto, E. Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc. Natl Acad. Sci. USA* **93**, 12845–12850 (1996).
- Wolffe, A. P. Histone deacetylase: a regulator of transcription. *Science* **272**, 408–411 (1996).
- Wade, P. A., Pruss, D. & Wolffe, A. P. Histone acetylation: chromatin in action. *Trends Biochem. Sci.* (in the press).
- Horlein, A. J. *et al.* Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature* **377**, 397–404 (1995).
- Zamir, I. *et al.* A nuclear hormone receptor corepressor mediates transcriptional silencing by receptors with distinct repression domains. *Mol. Cell. Biol.* **16**, 5458–5465 (1996).
- Vojtek, A. B., Hollenberg, S. M. & Cooper, J. A. Mammalian Ras interacts directly with the serine/threonine kinase Raf. *Cell* **74**, 205–214 (1993).
- Matallana, E., Franco, L. & Perez-Ortin, J. E. Chromatin structure of the yeast *SUC2* promoter in regulatory mutants. *Mol. Gen. Genet.* **231**, 395–400 (1992).
- Cooper, J. P., Roth, S. Y. & Simpson, R. T. The global transcriptional regulators, SSN6 and TUP1, play distinct roles in the establishment of a repressive chromatin structure. *Genes Dev.* **8**, 1400–1410 (1994).
- Tzamaras, D. & Struhl, K. Functional dissection of the yeast Cyc8-Tup1 transcriptional co-repressor complex. *Nature* **369**, 758–761 (1994).
- Herschbach, B. M., Arnaud, M. B. & Johnson, A. D. Transcriptional repression directed by the yeast alpha-2 protein *in vitro*. *Nature* **370**, 309–311 (1994).
- Wong, J., Shi, Y. B. & Wolffe, A. P. A role for nucleosome assembly in both silencing and activation of the *Xenopus* TR beta A gene by the thyroid hormone receptor. *Genes Dev.* **9**, 2696–2711 (1995).
- Khavari, P. A., Peterson, C. L., Tamkun, J. W. & Crabtree, G. R. BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. *Nature* **366**, 170–174 (1993).
- Muchardt, C. & Yaniv, M. A human homologue of *Saccharomyces cerevisiae* SNF2/SWI2 and *Drosophila* brm genes potentiates transcriptional activation by the glucocorticoid receptor. *EMBO J.* **12**, 4279–4290 (1993).
- Chiba, H., Muramatsu, M., Nomoto, A. & Kato, H. Two human homologues of *Saccharomyces cerevisiae* SWI2/SNF2 and *Drosophila* brm are transcriptional coactivators cooperating with the estrogen receptor and the retinoic acid receptor. *Nucleic Acids Res.* **22**, 1815–1820 (1994).
- Dunaief, J. L. *et al.* The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* **79**, 119–130 (1994).
- Strober, B. E., Dunaief, J. L., Guha, S. & Goff, S. P. Functional interactions between the hBRM/hBRG1 transcriptional activators and the pRB family of proteins. *Mol. Cell. Biol.* **16**, 1576–1583 (1996).
- Borrow, J. *et al.* The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nature Genet.* **14**, 33–41 (1996).
- Bannister, A. I. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641–643 (1996).
- Yang, X.-L., Ogryzko, V. V., Nishikawa, J., Howard, B. H. & Nakatani, Y. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**, 319–324 (1996).
- Haupt, Y., Alexander, W. S., Barri, G., Klinken, S. P. & Adams, J. M. Novel zinc finger gene implicated as myc collaborator by retrovirally accelerated lymphomagenesis in Eμ-myc transgenic mice. *Cell* **65**, 753–763 (1991).
- van Lohuizen, M. *et al.* Identification of cooperating oncogenes in Eμ-myc transgenic mice by provirus tagging. *Cell* **65**, 737–752 (1991).
- van Lohuizen, M., Frasch, M., Wientjens, E. & Berns, A. Sequence similarity between the mammalian bmi-1 proto-oncogene and the *Drosophila* regulatory genes *Psc* and *Su(2)2*. *Nature* **353**, 353–355 (1991).
- Evan, G. I. *et al.* Induction of apoptosis in fibroblasts by c-myc protein. *Cell* **69**, 119–128 (1992).
- Roy, A. L., Carruthers, C., Gutjahr, T. & Roder, R. G. Direct role for Myc in transcription initiation mediated by interactions with TFIIB. *Nature* **365**, 359–361 (1993).
- Li, L. H., Nerlov, C., Prendergast, G., MacGregor, D. & Ziff, E. B. c-Myc represses transcription *in vivo* by a novel mechanism dependent on the initiator element and Myc box II. *EMBO J.* **13**, 4070–4079 (1994).
- Lee, L. A., Dolde, C., Barret, J., Wu, C. S. & Dang, C. V. A link between c-Myc-mediated transcriptional repression and neoplastic transformation. *J. Clin. Invest.* **97**, 1687–1695 (1996).
- Galaktionov, K., Chen, X. & Beach, D. Cdc25 cell-cycle phosphatase as a target of c-myc. *Nature* **382**, 511–517 (1996).
- Wolffe, A. P. & Prusse, D. Targeting chromatin disruption: transcription regulators that acetylate histones. *Cell* **84**, 817–819 (1996).
- Brownell, J. E. & Allis, C. D. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr. Opin. Genet. Dev.* **6**, 176–184 (1996).
- Braunstein, M., Rose, A. B., Holmes, S. G., Allis, C. D. & Broach, J. R. Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev.* **7**, 592–604 (1993).
- Rundlett, S. E. *et al.* HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. *Proc. Natl Acad. Sci. USA* **93**, 14503–14508 (1996).
- De Rubertis, F. *et al.* The histone deacetylase RPD3 counteracts genomic silencing in *Drosophila* and yeast. *Nature* **384**, 589–591 (1996).
- DePinho, R. A., Schreiber-Agus, N. & Alt, F. W. Myc family oncogenes in the development of normal and neoplastic cells. *Adv. Cancer Res.* **57**, 1–46 (1991).
- Ayer, D. E., Kretzner, L. & Eisenman, R. N. Mad: a heterodimeric partner of Max that antagonizes Myc transcriptional activity. *Cell* **72**, 211–222 (1993).
- Yew, P. R., Liu, X. & Berk, A. J. Adenovirus E1B oncoprotein tethers a transcriptional repression domain to p53. *Genes Dev.* **8**, 190–202 (1994).

Acknowledgements. We thank A. Wolffe, A. Skoltschik, G. Yancopoulos, J. Brownell and members of the DePinho lab for critical reviews and comments; R. Sternglanz for advice regarding two-hybrid screens; P. Guida for assistance in the mSin3a bait characterization. This work was supported by grants from the NIH and an award from the Irma T. Hirsch Fund to R.A.D.; an NIH physician scientist and James S. McDonnell Foundation scholar award to L.A.; and NIH postdoctoral fellowship to N.S.A.; and an NCI Cancer core grant.

Correspondence and requests for materials should be addressed to R.A.D. (e-mail: depinho@acem.yu.edu).

An X-ray superflare from an infrared protostar

N. Grosso*, T. Montmerle*, E. D. Feigelson†, P. André*, S. Casanova‡ & J. Gregorio-Hetem§

* Service d'Astrophysique, CEA/DAPNIA/SAP, Centre d'études de Saclay, 91191 Gif-sur-Yvette Cedex, France

† Department of Astronomy & Astrophysics, Pennsylvania State University, University Park, Pennsylvania 16802, USA

‡ Observatoire Astronomique de Strasbourg, 11, rue de l'Université, 67000 Strasbourg, France

§ IAG/USP, São Paulo University, Brazil

Class I protostars¹ are very young, low-mass stellar objects that are, according to current models, composite: they include a central star (still in the process of formation) surrounded by an accretion disk $\sim 10\text{--}100$ AU in radius and embedded in an extended, infalling envelope of gas and dust up to $\sim 10^4$ AU in size². X-ray emission from such protostars has recently been reported^{3,4}, suggesting that X-ray ionization of gas and heating of dust could profoundly influence the physical and chemical properties of young stellar systems. But these observations did not have the resolution necessary to rule out a non-protostar origin for the emissions. Here we report X-ray observations of one of the nearest star-forming regions—the ρ Ophiuchi cloud—that clearly show an intense X-ray flare associated with a deeply embedded protostar. The peak X-ray luminosity, after correcting for extinction, is $\geq 10\text{--}100$ times the Sun's bolometric luminosity. The behaviour and intensity of the flare can be modelled as arising from a magnetically confined, low-density plasma bubble $\sim 0.05\text{--}0.3$ AU in diameter (much larger than the star itself), and the X-ray luminosity equals or exceeds the bolometric luminosity of the forming star. Taken together, the evidence suggests that the X-rays are not created by the type of magnetic activity seen on the Sun or on other young low-mass stars, but rather are associated with processes in the circumstellar accretion disk or within the envelope.

A deep ROSAT Position Sensitive Proportional Counter (PSPC) exposure of the ρ Ophiuchi cloud shows many X-ray sources clustered around two of the densest molecular cores (A and F), which are sites of current star formation⁵; seven sources coincide with class I protostars, but possible faint non-protostellar counterparts are also found in the same error boxes. A group of class I protostars is more clearly detected with the ASCA satellite in the Coronet cluster⁶, but the individual stars are only partially resolved. The definite confirmation of the detection of X-ray emission from class I protostars thus requires high-angular-resolution observations.

Using the ROSAT High Resolution Imager (HRI)⁷, we obtained a deep exposure of the ρ Oph core F. This exposure was performed in two observations totalling 40,013 s: the first was obtained between 9 and 14 March 1995 (12,473 s), and the second between 18 and 20 August 1995 (27,540 s); both were centred on the same position (right ascension (RA) 16 h 27 min 26.4 s, declination (dec.) $-24^\circ 40' 48''$; J2000 coordinates), in the middle of a group of class I sources (Fig. 1). After astrometric correction (Fig. 2), source 6 (RA 16 h 27 min 26.9 s, dec. $-24^\circ 40' 51''$ (J2000) with an accuracy of $\text{rms}_{\text{total}} = 2.6''$; $\text{rms}_{\text{total}}$ is defined in Fig. 2), is found to be located only 0.4 $\text{rms}_{\text{total}}$ away from the class I source YLW15⁸, and 2.5 $\text{rms}_{\text{total}}$ away from another infrared source, GY263. The identification of source 6 with YLW15 is thus very reliable, especially as this source lies almost at the centre of the HRI field, where the best angular resolution and sensitivity are obtained.

YLW15, or IRS43, is a well-known source of the ρ Oph cloud, studied over the entire near-infrared⁹ to millimetre (ref. 10) range.

The slope of its infrared spectral energy distribution (SED), $\alpha_{\text{IR}} = d(\lambda F_\lambda)/d\lambda = 1.2$, led to its initial classification as a class I protostar. Its bolometric luminosity estimated by the calorimetric method⁷, including the recent millimetre measurements¹⁰, is $L_{\text{bol}} = (10 \pm 1)L_\odot$ (where $L_\odot$ is the solar luminosity), making it one of the most luminous protostars in the core F region. Using the ratio of the millimetric luminosity to the bolometric luminosity as an evolutionary indicator¹¹, YLW15 appears to be one of the youngest self-embedded infrared sources of the ρ Oph cluster¹². In addition, with a 6-cm flux density of 3.3 mJy, it is by far the brightest stellar radio continuum source of core F⁶: very-long-baseline interferometry (VLBI) measurements¹³ suggest the radio flux is probably due to free-free emission of an ionized jet at the base of the recently discovered molecular outflow¹². With all these features, YLW15 is confirmed to be a bona fide class I protostar, at an intermediate stage of evolution between class 0 protostars¹¹ and 'flat-spectrum' T Tauri stars. The estimated age of such protostars is $10^5\text{--}10^6$ yr. In particular, YLW15 is probably significantly younger than the flat spectrum ($\alpha_{\text{IR}} \approx 0$) sources of the Coronet cluster⁴, with an estimated age of at least a few 10^5 yr. YLW15 may thus be representative of the youngest low-mass stars detectable in X-rays.

YLW15 displayed a strong flare during our X-ray observations. During the second orbit of the first observation, the HRI detected a count rate of 16.7 ± 5 counts per ks, which subsequently declined to levels consistent with non-detection (Fig. 3). It is clear that YLW15 could be detected only because it flared. ROSAT detected 12 photons (in 12 min) at maximum, and 39 photons during the whole flare observation (2 h 10 min). During the event, the X-ray flux increased by at least a factor of 20, then decayed with an e -folding time $\tau_d \approx 5$ hours.

We interpret such a time behaviour as the post-flare cooling of a hot plasma: estimating the physical conditions of this plasma will help determine the origin of the flare. We use a simple model where a magnetically confined sphere of uniform density emits via thermal bremsstrahlung, including line emission (a Raymond-Smith spectrum). We assume that the energy is supplied in a time $\ll \tau_d$ at the beginning of the flare, and that the plasma subsequently cools radiatively. The corresponding cooling timescale is $\tau_r = 3kT_X/n_e P(T_X)$, where k is the Boltzmann constant, n_e is the electron density, and $P(T_X)$ is the radiative cooling function¹⁴, which we fit by $P(T_X) = AT_X^\alpha$ (in $\text{erg cm}^3 \text{s}^{-1}$): for $T_X \approx (1\text{--}2) \times 10^7$ K, $A = 2.88 \times 10^{-19}$ and $\alpha = -0.6$, and $T_X > 2 \times 10^7$ K, $A = 1.86 \times 10^{-25}$ and $\alpha = +0.25$. The average electron densities in the emitting volume are then obtained by setting $\tau_r = \tau_d$. Unfortunately, T_X cannot be measured because the HRI has no spectral resolution. We thus explore a wide range of temperatures, considering that $T_X \approx (1\text{--}4) \times 10^7$ K is typical of flaring T Tauri stars¹⁵, and $T_X \approx (7\text{--}8) \times 10^7$ K has been found in the Coronet cluster⁴. For $T_X = 10^7\text{--}10^8$ K, we find $n_e \approx (2\text{--}10) \times 10^{10} \text{ cm}^{-3}$, and, from the emission measures, plasma diameters $l \approx (10\text{--}60)R_\odot = 0.05\text{--}0.3$ AU (with a visual extinction $A_V = 20\text{--}40$ mag, see below; $R_\odot$ is the solar radius). Assuming equipartition between the magnetic pressure $B^2/8\pi$ and the ionized gas pressure $2n_e kT_X$, the magnetic field confining the sphere is at least $B = 40\text{--}300$ G. Details are given in Table 1.

The determination of the X-ray luminosity of the YLW15 flare depends critically on the extinction suffered by the X-rays due to the circumstellar envelope and to the interstellar medium along the line of sight. Millimetre continuum observations indicate that YLW15 is surrounded by a relatively dense, compact dusty envelope with average density $n_{\text{H}_2} \approx 10^5 \text{ cm}^{-3}$, outer radius $R_{\text{out}} \leq 3,000$ AU, and total (gas + dust) mass $M_{\text{env}} = (0.05\text{--}0.3)M_\odot$ (ref. 10, and F. Motte, P.A. and R. Neri, manuscript in preparation). A recent large-scale 1.3-mm bolometer-array map taken at the IRAM 30-m telescope shows that the envelope of YLW15 is deeply embedded within, and smoothly blends into, the ρ Oph F dense core (F. Motte *et al.*, manuscript in preparation). These measurements yield, for the

envelope alone, a column density $N_{\text{H,env}} = (2-9) \times 10^{22} \text{ cm}^{-2}$ (depending on the details of the merging between the envelope and the dense core), and a visual absorption $A_{\text{V,env}} = 10-40$ (using the recently updated conversion value $N_{\text{H}} = 2.23 \times 10^{21} A_{\text{V}} \text{ cm}^{-2} \text{ mag}^{-1}$; ref. 16). They also suggest that YLW15 is embedded in a cloud core, so that the total extinction $A_{\text{V,tot}} = A_{\text{V,env}} + A_{\text{V,cloud}}$ must be higher by several magnitudes at least. An estimate of $A_{\text{V,tot}}$ may be derived from $A_{\text{V,tot}} \approx 12[(H-K) - (H-K)_0]$ (ref. 7), where $(H-K)_0$ is the intrinsic near-infrared colour index of the central protostar and circumstellar disk. With $H = 13.61$, $K = 9.96$ (ref. 7), we obtain $A_{\text{V,tot}} \approx 44$ if $(H-K)_0 \approx 0$ (no disk) and $A_{\text{V,tot}} \approx 33$ if $(H-K)_0 \approx 0.9$ (disk with flat spectrum). Another estimate of $A_{\text{V,tot}}$ can be obtained by modeling the near-infrared portion of the SED as in Myers *et al.*¹⁷. Taking published values of $L_{\text{bol,*}}$ and of J, H, K, L

photometry⁷, we find $A_{\text{V,tot}} > 31 \pm 2$. We conclude that $A_{\text{V,tot}}$ is very likely to be in the 20–40 magnitude range, with a best estimate $A_{\text{V,tot}} \approx 30$. A future improvement may come from more direct measurements of A_{V} based on near-infrared imaging polarimetry¹⁸.

With a distance of 160 pc, an isothermal Raymond–Smith plasma spectrum, and standard solar elemental abundances, we use the *PROS* and *XSPEC* software packages to determine the X-ray luminosities corresponding to the HRI flux of 0.0167 counts per s observed at the peak of the flare (Table 1). With $A_{\text{V}} \approx 20-40$, we find a total intrinsic X-ray luminosity (that is integrated over all X-ray energies) $L_{\text{X,tot}} \approx 10^{34}-10^{36} \text{ erg s}^{-1}$, and, defining an ‘X-ray efficiency factor’ $\eta_{\text{X}} = L_{\text{X,tot}}/L_{\text{bol,*}}$, $\eta_{\text{X}} \geq 1-100$. Both values are much larger than found so far for flares on young stars. For example, the very luminous ($L_{\text{bol,*}} = 33L_{\odot}$) Orion weak-lined T Tauri star P1724

Table 1 Model plasma parameters and luminosities at flare peak

	$T_{\text{X}} (10^7 \text{ K})$	1	3	6	12
	$n_{\text{e}} (10^{10} \text{ cm}^{-3})$	1.6	5.4	8.0	13
	B (gauss)	36	110	180	330
$A_{\text{V}} = 20$	$\log L_{\text{X,HRI}}^{\text{RS}} (10^{31} \text{ erg s}^{-1})$	31.3	31.5	31.5	31.5
	$\log L_{\text{X,HRI}}^{\text{ext}} (10^{31} \text{ erg s}^{-1})$	33.2	32.9	32.9	32.9
	$\log L_{\text{X,tot}} (10^{31} \text{ erg s}^{-1})$	33.7	34.0	34.3	34.5
	$l/R_{\odot}$	17	13	12	11
$A_{\text{V}} = 30$	$\log L_{\text{X,HRI}}^{\text{RS}} (10^{31} \text{ erg s}^{-1})$	31.6	31.7	31.7	31.8
	$\log L_{\text{X,HRI}}^{\text{ext}} (10^{31} \text{ erg s}^{-1})$	33.8	33.6	33.5	33.5
	$\log L_{\text{X,tot}} (10^{31} \text{ erg s}^{-1})$	34.6	35.1	35.2	35.5
	$l/R_{\odot}$	35	27	25	23
$A_{\text{V}} = 40$	$\log L_{\text{X,HRI}}^{\text{RS}} (10^{31} \text{ erg s}^{-1})$	31.8	32.0	32.0	32.0
	$\log L_{\text{X,HRI}}^{\text{ext}} (10^{31} \text{ erg s}^{-1})$	34.0	34.1	34.1	34.1
	$\log L_{\text{X,tot}} (10^{31} \text{ erg s}^{-1})$	35.0	35.9	36.1	36.4
	$l/R_{\odot}$	65	52	49	44

The top three rows give the model plasma parameters at the observed flare peak, assuming a magnetically confined plasma sphere. Because the HRI has no spectral resolution, we have to explore a wide range of temperatures T_{X} (see text). n_{e} is the electronic density, assumed to be uniform, B is the minimum value of the confining magnetic field, and l is the sphere diameter. The lower rows give the peak flare luminosities, calculated assuming a Raymond–Smith spectrum, as a function of the estimated range of visual extinction (see text). $L_{\text{X,HRI}}^{\text{RS}}$ is the apparent luminosity in the HRI bandwidth; $L_{\text{X,HRI}}^{\text{ext}}$ is the same, corrected for extinction; $L_{\text{X,tot}}$ is the total intrinsic luminosity, integrated over all X-ray energies after correction for bandwidth and extinction. These values are actually lower limits, as the flare maximum may have occurred before our observation (see Fig. 3). The software packages (*PROS* and *XSPEC*) used to calculate X-ray luminosities are available at <http://hea-www.harvard.edu/PROS/pros.html> and http://legacy.gsfc.nasa.gov/docs/xanadu/xspec/u_manual.html, respectively.

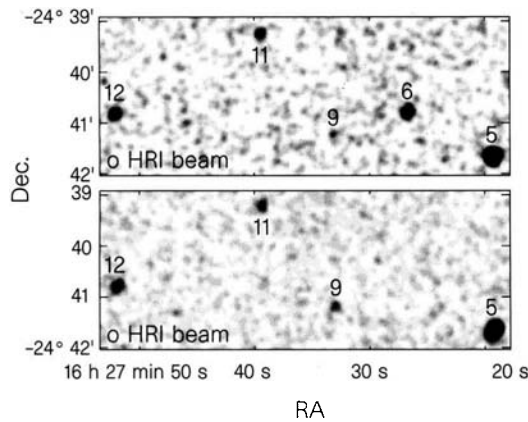


Figure 1 Detail of the HRI X-ray images of the ρ Oph core F. The HRI angular resolution is $5''$ full-width at half-maximum on-axis (HRI beam), compared to $25''$ for the PSPC. It is sensitive to the 0.1–2.4 keV energy range, but has no spectral resolution. To increase signal-to-noise ratio (S/N) we remove pulse height analyser channels 9–16 owing to high noise levels. Raw pixels of $0.5'' \times 0.5''$ were combined to form pixels of $2'' \times 2''$, and smoothed with a gaussian ($\sigma = 3''$). Using a S/N criterion, with background noise level estimated from a large source-free region, we find 13 sources with $S/N > 3$ in the whole field (N.G. *et al.*, manuscript in preparation). The enlarged images shown are normalized to identical exposures (12,473 s). **a**, First observing period between 9 and 14 March 1995 (12,473 s); **b**, second observing period between 18 and 20 August 1995 (27,540 s). The variability of source 6 is clearly visible, in contrast to the unchanged intensity of the other sources. Sources 5 and 12 have unambiguous optically visible stellar counterparts, respectively SR12AB/ROXs21 and IRS55/ROXs31. Right ascension (RA) and declination (dec.) are given in J2000 coordinates.

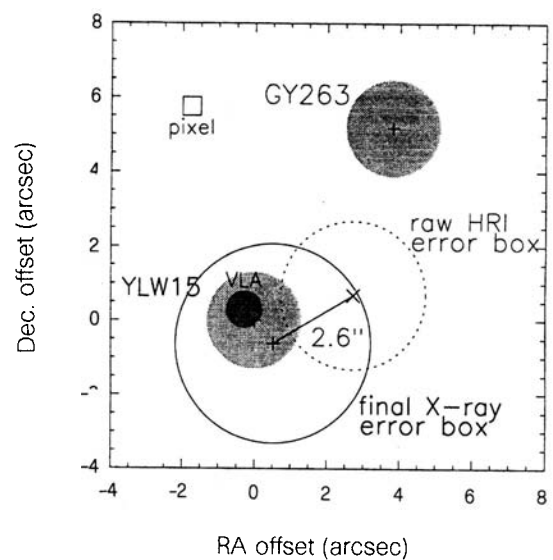


Figure 2 Relative positions and error boxes of YLW15, GY 263 and HRI source 6 before and after boresight error correction. The HRI $0.5'' \times 0.5''$ pixel size is also shown. The r.m.s. error position of GY 263 is $1.3''$ and $0.5''$ for the VLA position of YLW15⁶. The raw HRI error box is $\text{rms}_{\text{source}} = 2''$. Comparison of the X-ray positions with unambiguous optically visible stellar counterparts of six sources in the whole HRI field reveals an average offset of $2.6''$. After correction for this satellite boresight error, X-ray and optical positions agree within $\text{rms}_{\text{shift}} = 1.8''$, consistent with the HRI position accuracy. Thus, the HRI error box after boresight error correction is $2.6''$ ($\text{rms}_{\text{total}}^2 = \text{rms}_{\text{source}}^2 + \text{rms}_{\text{shift}}^2$). Source 6 is found $1.2'' = 0.4 \text{ rms}_{\text{total}}$ away from YLW15, and $6.7'' = 2.5 \text{ rms}_{\text{total}}$ from GY 263⁶ (coordinate offsets are with respect to the YLW15 position published in ref. 8).

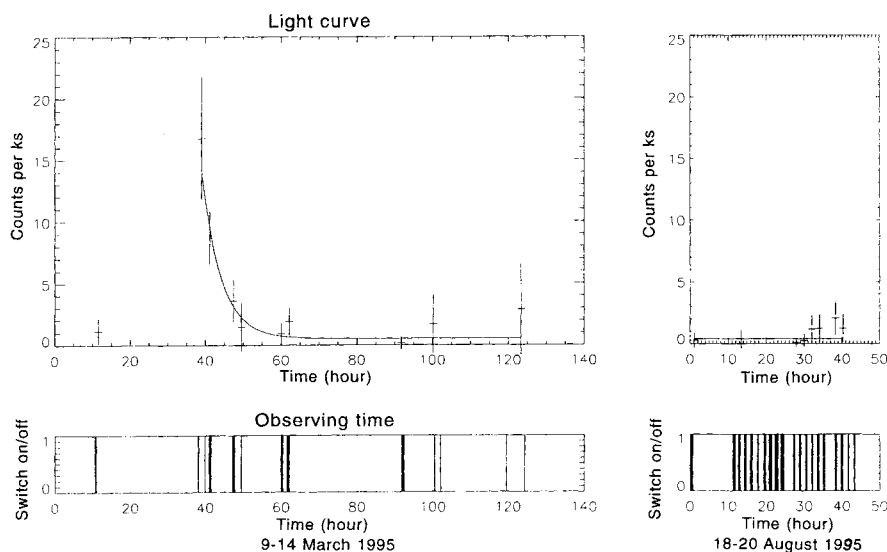


Figure 3 Light curve of YLW15 using 2-hour bins and 1σ error bars, deduced from observations of various ROSAT orbits between 9 and 14 March 1995 (12,473 s), and between 18 and 20 August 1995 (27,540 s). The 2-hour bin count rate was estimated from photon counts during the observing sessions (vertical bars in the lower panel). The continuous line in the first exposure is a least-square fit to an exponential decay, and weights equal to $1/\sigma^2$: $N(t) = 0.5 \pm 0.4 + (13.7 \pm 3.4) \times \exp[-t/(5.1 \pm 1.9 \text{ hours})]$ counts per ks. The quiescent level is 0.41 ± 0.47 counts per ks, consistent in both observations with the end of the exponential decay.

exhibited a flare with peak X-ray luminosity $L_X \approx 1.8 \times 10^{33} \text{ erg s}^{-1}$ (ref. 19), thus $\eta_X \approx 0.01$; the classical T Tauri star LkH92 ($L_{\text{bol},*} = 4L_\odot$) had peak $L_X \approx 5 \times 10^{32} \text{ erg s}^{-1}$ (ref. 20) and $\eta_X \approx 0.03$.

The remarkably high value of η_X and the minimum size of the X-ray emitting region l (Table 1) for YLW15 require a flare much more powerful and larger than seen in any stellar analogue of solar-type magnetic activity. The size of the magnetic structure producing this flare, particularly if elongated rather than spherical geometries are considered, is comparable to the co-rotation radius of a protostellar circumstellar disk or the size of its inner envelope cavity. We are thus compelled to consider non-solar-type origins for the flare where the magnetic structure is connected to circumstellar components, while maintaining the underlying mechanism of explosive release of energy stored in magnetic fields through reconnection, that is, interaction of magnetic fluxes having antiparallel components. In protostars, magnetic field reconnection can take place during protostellar infall²¹, or later as a result of shearing between the stellar magnetosphere and the circumstellar disk at several locations: at the magnetosphere-disk Alfvén surface^{22,23} or co-rotation radius²⁴, at the X-point above the co-rotation radius where magnetic geometries change from closed to open²⁵, and at the interface between a stellar wind and the disk²⁶. Another possibility of reconnection arises if the central object in YLW15 is a double system with magnetic fields lines linking the two components, in a fashion reminiscent of RS CVn binaries, but with much larger separation.

Irrespective of the exact location where the X-rays are produced in the protostellar system, our result shows that large numbers of hard photons are likely to be present. They will efficiently photo-ionize the circumstellar envelopes and, if source geometries are favourable, the protostellar accretion disk, and also affect their chemistry (compare refs 3, 27). In particular, protostellar X-rays may regulate the rate of accretion from the disk onto the star²⁸. The disk would also be bombarded by shocks and energetic protons associated with the X-ray flares. As the early Sun should have passed through the flaring class I protostellar phase, the existence of such shocks, as well as X-ray and particle irradiation phenomena, may resolve several long-standing mysteries (isotopic anomalies, chondrule melting, excess particle irradiation) in ancient meteorites (refs 26, 29–31). □

Received 22 January; accepted 25 March 1997.

1. Lada, C. J. in *Star Forming Regions* (eds Peimbert, M. & Jugaku, J.) 1–18 (Proc. 115th IAU Symp., Kluwer, Dordrecht, 1987).
2. Shu, F. H., Adams, F. C. & Lizano, S. *Annu. Rev. Astron. Astrophys.* **25**, 23–81 (1987).
3. Casanova, S., Montmerle, T., Feigelson, E. D. & André, P. *Astrophys. J.* **439**, 752–770 (1995).
4. Koyama, K., Ueno, S., Kobayashi, N. & Feigelson, E. *Publ. Astron. Soc. Jpn* **48**, 87–92 (1996).

5. David, L. P., Harnden, F. R., Kearns, K. E. & Zombeck, M. V. *The ROSAT HRI* (US ROSAT Science Data Center/SAO, NASA/GSFC, Greenbelt, DC, 1996).
6. Leous, J. A., Feigelson, E. D., André, P. & Montmerle, T. *Astrophys. J.* **379**, 683–688 (1991).
7. Wilking, B. A., Lada, C. J. & Young, E. T. *Astrophys. J.* **340**, 823–852 (1989).
8. Greene, T. P. & Young, E. T. *Astrophys. J.* **395**, 516–528 (1992).
9. Strom, K., Kepner, J. & Strom, S. E. *Astrophys. J.* **438**, 813–829 (1995).
10. André, P. & Montmerle, T. *Astrophys. J.* **420**, 837–862 (1994).
11. André, P., Ward-Thompson, D. & Barsony, M. *Astrophys. J.* **406**, 122–141 (1993).
12. Bontemps, S., André, P., Terebey, S. & Cabrit, S. *Astron. Astrophys.* **311**, 858–872 (1996).
13. André, P., Deeney, B. D., Phillips, R. B. & Lestrade, J. F. *Astrophys. J.* **401**, 667–677 (1992).
14. Van den Oord, G. H. J. & Mewe, R. *Astron. Astrophys.* **213**, 245–260 (1989).
15. Montmerle, T. & Casanova, S. in *Magnetodynamic Phenomena in the Solar Atmosphere—Prototypes of Stellar Magnetic Activity* (eds Uchida, Y. et al.) 247–258 (Kluwer, Dordrecht, 1996).
16. Ryter, C. E. *Astron. Space Sci.* **236**, 285–291 (1996).
17. Myers, P. C. et al. *Astrophys. J.* **319**, 340–357 (1987).
18. Whitney, B. A. & Kenyon, S. J. *Astrophys. J.* (in the press).
19. Preibisch, T., Neuhauser, R. & Alcalá, J. M. *Astron. Astrophys.* **304**, 13–16 (1995).
20. Preibisch, T., Zinnecker, H. & Schmitt, J. H. M. M. *Astron. Astrophys.* **279**, 33–36 (1993).
21. Norman, C. & Heyvaerts, J. *Astron. Astrophys.* **147**, 247–256 (1985).
22. Hayashi, M. R., Shibata, K. & Matsumoto, R. *Astrophys. J.* **468**, L37–L40 (1996).
23. Aly, J. J. & Kuipers, J. *Astron. Astrophys.* **227**, 473–482 (1990).
24. Königl, A. *Astrophys. J.* **370**, L39–L43 (1991).
25. Najita, J. R. & Shu, F. H. *Astrophys. J.* **429**, 808–825 (1994).
26. Cameron, A. G. W. *Meteoritics* **30**, 133–161 (1995).
27. Maloney, P. R., Hollenbach, D. J. & Tielens, A. G. G. M. *Astrophys. J.* **466**, 561–584 (1996).
28. Glassgold, A. E., Najita, J. & Igea, J. *Astrophys. J.* (in the press).
29. Feigelson, E. D. *Icarus* **51**, 155–163 (1982).
30. Woolum, D. S. & Hohenberg, C. in *Protostars & Planets III* (eds Levy, E. H. & Lunine, J. H.) 903–919 (Univ. Arizona Press, Tucson, 1993).
31. Lee, T., Shu, F. H., Shang, H. & Glassgold, A. E., *Science* (submitted).

Correspondence should be addressed to T.M. (e-mail: montmerle@sapvsg.saclay cea.fr).

Stable knot-like structures in classical field theory

L. Faddeev*† & Antti J. Niemi‡†

* St Petersburg Branch of Steklov Mathematical Institute, Russian Academy of Sciences, Fontanka 27, St Petersburg, Russia

† Helsinki Institute of Physics, PO Box 9, FIN-00014 University of Helsinki, Finland

‡ Department of Theoretical Physics, Uppsala University, PO Box 803, S-75108 Uppsala, Sweden

In 1867, Lord Kelvin proposed that atoms—then considered to be elementary particles—could be described as knotted vortex tubes in ether¹. For almost two decades, this idea motivated an extensive study of the mathematical properties of knots, and the results obtained at that time by Tait² remain central to mathematical knot theory^{3,4}. But despite the clear relevance of knots to a large number of physical, chemical and biological systems, the physical

properties of knot-like structures have not been much investigated. This is largely due to the absence of a theoretical means for generating stable knots in the nonlinear field equations that can be used to describe such systems. Here we show that knot-like structures can emerge as stable, finite-energy solutions in one such class of equations—local, three-dimensional lagrangian field-theory models. Our results point to several experimental and theoretical situations where such structures may be relevant, ranging from defects in liquid crystals and vortices in superfluid helium to the structure-forming role of cosmic strings in the early Universe.

Today it is commonly accepted that fundamental interactions are described by string-like structures⁵, with different elementary particles corresponding to the vibrational excitations of a fundamental string. But even though there are hints of connections between modern string theory and knot theory⁶, knot-like structures have not yet been of much significance. However, knots do already enter in a number of other physical scenarios. These include models in statistical physics and hydrodynamics, quantum chromodynamics, cosmic strings that are expected to be responsible for structure formation in the early Universe, and approaches to quantum gravity where knots are supposed to determine gauge-invariant observables. String-like vortices appear in type-II superconductors where they confine magnetic fields within the cores of vortex-like structures, and similar phenomena are also present in superfluid ⁴He. Recent experiments with nematic liquid crystals⁷ and ³He superfluids^{8,9} have also revealed interesting vortex structures that can be described by theoretical methods which are adopted from cosmic-string models. The physics of knots is also rapidly becoming an important issue in polymer physics, and in particular in molecular biology where entanglement of a DNA chain interferes with vital processes of replication, transcription and recombination¹⁰.

To investigate the properties of knot-like structures from first principles, one needs a theoretical model where knots emerge as solitons, that is, as stable finite energy solutions to the pertinent nonlinear field equations. The literature on solitons is enormous, and there are several extensive reviews^{11–13}. However, until now the activity has mainly concentrated on one dimension with the notable exceptions of the two-dimensional vortex and nonlinear σ -model soliton, and skyrmions and 'tHooft–Polyakov monopoles in three dimensions. These are all point-like configurations, and can not be

directly associated with knot-like structures.

When embedded in three dimensions, a point-like two-dimensional soliton becomes a line vortex. For a finite energy, its length must be finite which is possible if its core forms a knot. In 1975 one of us proposed^{14,15} that closed, knotted vortices could be constructed in a definite dynamical model. The explicit solution conjectured in refs 14, 15 has the shape of a doughnut; it is a closed torus-like vortex ring, twisted once around its core before joining the ends to ensure stability against shrinking. In knot theory this closed vortex corresponds to the unknot, the simplest possible knot-like structure. However, despite numerous attempts no such stable configuration has been constructed, neither by explicit analysis nor by a numerical investigation.

We have employed numerical algorithms in powerful computers to investigate the model introduced in ref. 14 and 15, and we have found strong evidence for the existence of the doughnut-shaped unknot vortex. In addition we have found indications supporting the existence of a trefoil vortex (see Fig. 4). A trefoil is the simplest non-trivial example of a torus knot. This is a general family of knots which can be classified by counting how many times a knot winds a torus, both around and along the direction of the doughnut^{3,4}. Our results then suggest that in fact all torus knots should appear as a vortex solitons in the model proposed in refs 14 and 15. This model describes the 3 + 1 dimensional dynamics of a three-component vector $\mathbf{n}(\mathbf{x}, \tau)$ with unit length, $\mathbf{n} \cdot \mathbf{n} = 1$. Such a vector field is a typical degree of freedom in the nonlinear σ -model, a prototype relativistic quantum field theory. It also appears as an order parameter in the Heisenberg ferromagnet. A unit vector field is also present in models of nematic liquid crystal where it characterizes the average direction of the rod, and in ³He-A superfluid where it determines the spin projection direction for a Cooper pair. Indeed, the model proposed in refs 14, 15 is quite universal, and we expect our results to have a large number of applications.

In order that $\mathbf{n}(\mathbf{x})$ describes a localized stationary knot, it must go to a constant vector $\mathbf{n}(\mathbf{x}) \rightarrow \mathbf{n}_0$ at large distances. Consequently $\mathbf{n}(\mathbf{x})$ defines a mapping from the compactified $R^3 \sim S^3 \rightarrow S^2$. Such mappings fall into nontrivial homotopy classes $\pi_3(S^2) \approx \mathbb{Z}$ and can be characterized by a topological invariant called the Hopf invariant^{3,4}, an integral formula which counts the number of times the two-sphere S^2 is covered by the three-sphere S^3 . For this we introduce the two-form $F = (d\mathbf{n} \wedge d\mathbf{n}, \mathbf{n})$ on the target S^2 . Its pre-image F_* on S^3 is

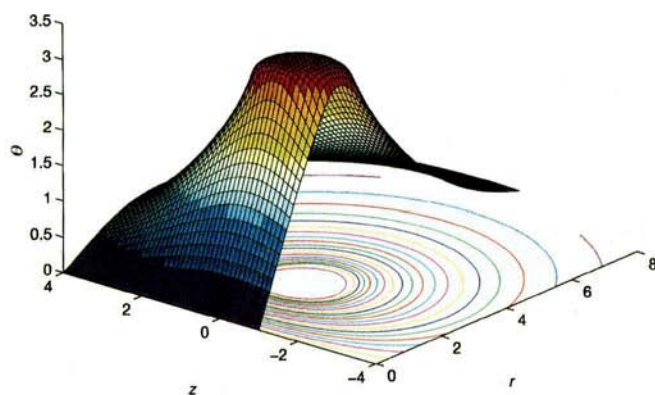


Figure 1 A combined cross-sectional surface and contour plot of $\theta(r, z)$ for a doughnut-shaped unknot vortex with $g^*/e \approx 0.24$. We use cylindrical coordinates (r, ϕ, z) where z coincides with the symmetry axis of the doughnut and r measures distance from the symmetry axis. The configuration is rotation-invariant around the z -axis, that is, along the doughnut. These rotations are parametrized by the angle ϕ which also labels the different cross-sectional planes. The configuration is $z \rightarrow -z$ symmetric.

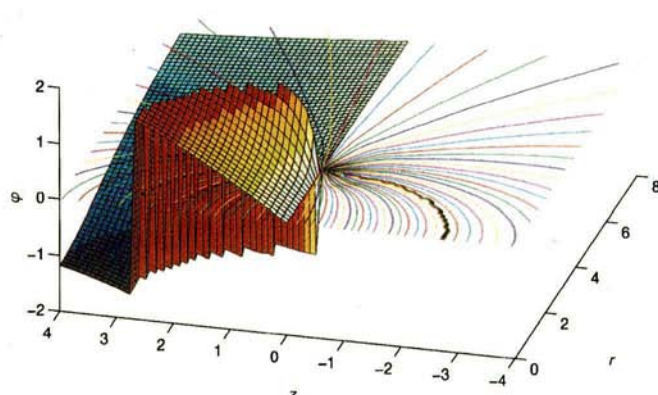


Figure 2 A combined cross-sectional surface and contour plot of $\phi(r, z)$ for the unknot vortex, in cylindrical coordinates as in Fig. 1. The line where ϕ jumps by π is clearly visible. Notice that due to the $z \rightarrow -z$ symmetry, ϕ changes by 2π when we go once around the core

exact, $F_* = dA_*$, and the Hopf invariant Q_H is given by

$$Q_H = \int_{R^3} F \wedge A \quad (1)$$

The hamiltonian proposed in refs 14 and 15 is

$$H = E_2 + E_4 = \int d^3x (g^2 (\partial_\mu \mathbf{n})^2 + e^2 F^2) \quad (2)$$

where g and e are coupling constants and $\mu = 1, 2, 3$. This is the most general three-dimensional hamiltonian that admits a relativistically invariant dynamics and involves only terms with no more than four derivatives. If we restrict to a sphere $S^2 \in SU(2)$, we can relate H to the $SU(2)$ Skyrme model that describes low-energy nuclear dynamics, but its topological features are different. In particular, the existence of non-trivial knotted vortex solutions in equation (2) is strongly suggested by the lower bound $H \geq c |Q_H|^{3/4}$, where the numerical constant c is known to be non-vanishing¹⁶.

In equation (2), the first term E_2 determines the standard non-linear $O(3)$ σ -model that describes the Heisenberg ferromagnet and admits static solitons in two dimensions. But in three dimensions a scaling $\mathbf{x} \rightarrow \rho \mathbf{x}$ reveals that stable stationary solutions are possible only if E_4 is also present. Indeed, under this scaling $E_2 \rightarrow \rho E_2$ but $E_4 \rightarrow \rho^{-1} E_4$ from which we conclude that in three dimensions finite-energy solutions obey the virial theorem,

$$E_2 = E_4 \quad (3)$$

The general properties of the unknot vortex in the model equation (2) have been analysed¹³. However, to our knowledge there have been no real attempts to find an actual solution. This is due to the fact, that even in the simplest case of an unknot, the separation of variables eliminates only one of the three space coordinates, leaving numerical methods as the sole alternative for finding a solution.

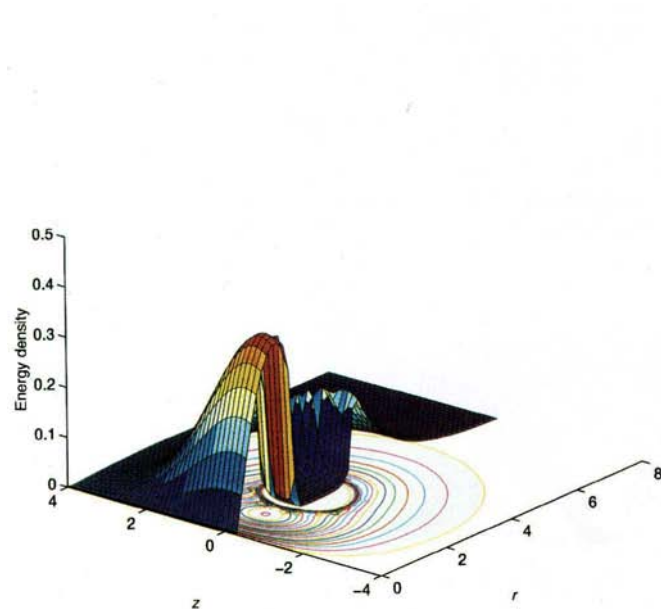


Figure 3 A combined surface and contour plot of energy density $H(r, z)$ for the unknot vortex, in cylindrical coordinates as in Fig. 1. The configuration is again $z \rightarrow -z$ symmetric. Our numerical simulations have revealed that for knots in the model equation (2) the energy density is very small (essentially vanishing) in a tube-like neighbourhood around the core, and admits a local maximum in a tyre-like neighbourhood around it. Consequently the distribution of energy density in this figure is generic to knots in the model equation (2), when viewed on their cross-sectional planes.

With the recent, enormous progress in supercomputing techniques serious attempts are finally becoming realistic.

The Euler-Lagrange equations for the hamiltonian in equation (2) determine a highly nonlinear elliptic system, and a direct numerical approach appears to be complicated. Instead, we consider the parabolic equation

$$\frac{d\phi_a}{dt} = - \frac{\delta H(\phi)}{\delta \phi_a} \quad (4)$$

where ϕ_a denotes a generic independent degree of freedom of the hamiltonian H in equation (2). The t -bounded solutions of equation (4) connect the critical points of H where the right-hand side of equation (4) vanishes, by flowing away from an unstable critical point at $t \rightarrow -\infty$ towards a stable critical point at $t \rightarrow +\infty$. From this we expect that by defining a suitable initial configuration at $t = 0$, for large $t \gg 0$ we flow towards a stable vortex solution of the original stationary equation.

We parametrize our unit vector by $\mathbf{n} = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta)$ through stereographic coordinates $\varphi = -\arctan(V/U)$ and $\theta = 2 \arctan \sqrt{U^2 + V^2}$ corresponding to the variables ϕ_a in equation (4). In terms of these variables the hamiltonian, equation (2), becomes

$$H = \int d^3x \frac{4g^2}{(1 + U^2 + V^2)^2} (\partial_\mu U^2 + \partial_\mu V^2) + \frac{16e^2}{(1 + U^2 + V^2)^4} (\partial_\mu U \partial_\nu V - \partial_\nu U \partial_\mu V)^2 \quad (5)$$

where $\mu, \nu = 1, 2, 3$. By a global $SO(3)$ rotation we can select our asymptotic vacuum vector $\mathbf{n}_0$ so that outside the vortex it points to the positive- z direction. This means that outside the vortex we are

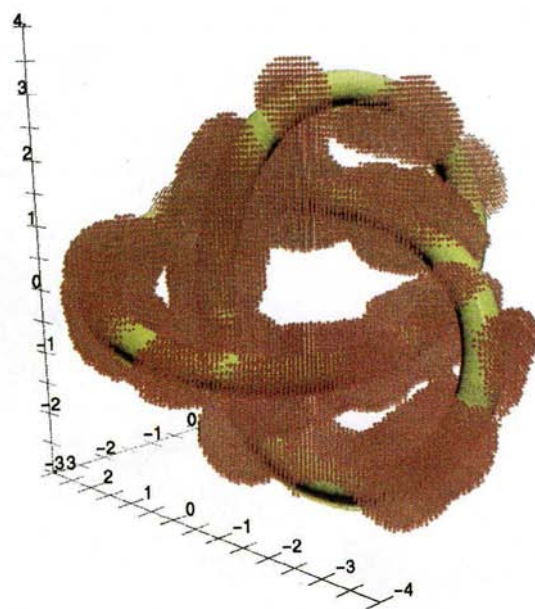


Figure 4 Three-dimensional spatial view of the extended core for the trefoil vortex with $g^2/e \approx 1.3$. A cross-section of the knot reveals an energy density distribution which is qualitatively similar to that in Fig. 3. The extended core corresponds to the region in Fig. 3 inside the knot, where energy density essentially vanishes. This region is described by the pink dots which have been computed from our finite-element approximation. The non-uniform pattern of the dots reflects the underlying lattice structure. The yellow centre denotes the core of the initial configuration which is determined using a minimum energy principle¹⁸. We thank J. Simon for providing us with the initial parametrization.

near the north pole where $\theta(\mathbf{x}) \approx 0$, while the core corresponds to the south pole $\theta(\mathbf{x}_c) = \pi$. The internal structure of a vortex can then be investigated by cutting it once with a plane at right angles to its core. This cross-sectional plane is topologically identical to a sphere S^2 : at the core we have $\theta(\mathbf{x}_c) = \pi$ corresponding to the south pole and outside the vortex on the plane we have $\theta(\mathbf{x}) \approx 0$ corresponding to the north pole. Furthermore, $\varphi(\mathbf{x})$ increases (or decreases, depending on orientation) by 2π when we go around the core once on the cross-sectional plane. (More generally, φ is defined modulo $2\pi n$ where n is an integer.) For U and V this means that we identify them as local coordinates on the Riemann sphere in a patch that contains the north pole $U(\mathbf{x}) = V(\mathbf{x}) = 0$. In Figs 1 and 2 we have drawn $\theta(\mathbf{x})$ and $\varphi(\mathbf{x})$ respectively for our unknot vortex. The general structure is clearly visible.

To specify the initial condition in equation (4) we need a knot configuration which is topologically equivalent to the desired vortex solution. We specify this configuration by first introducing a parametrization $\mathbf{x}(\lambda)$ for the centre of the knot in R^3 , and then for each λ we define local coordinates $U(\mathbf{x}; \lambda)$ and $V(\mathbf{x}; \lambda)$ on the cross-sectional planes. The choice of initial parametrization $\mathbf{x}(\lambda)$ for the core introduces a scale which may be quite different from the one specified by the coupling constant g in equation (5). To enhance convergence, we adopt a simple renormalization procedure by promoting g to a time-dependent function $g(t)$. We define this time dependence by demanding that at each value of t the virial theorem (equation (3)) must be obeyed. As a vortex solution obeys equation (3), this means that $g(t)$ flows toward a fixed point value g^* .

We have performed our numerical simulations using version 5.3 of the PDE2D finite-element algorithm¹⁷. During the early phase of our simulations, we used the initial knot configuration to determine the boundary conditions on the finite-element mesh. However, because the exact boundary conditions for the desired vortex are *a priori* unknown, after several time steps we decreased the size of the finite-element mesh and used the pertinent simulated configuration to determine the boundary conditions in the new mesh. By starting from a sufficiently large initial mesh, we then eventually obtain a submesh with boundary conditions that are close to those of the actual solution. For our simulation we used two Silicon Graphics Power Challenge computers with R8000 processors; one computer was equipped with 1 GB RAM, and the other with 2 GB.

For the unknot vortex ($Q_H = \pm 1$), the equations of motion can be simplified using axial symmetry. We select the symmetry axis of the doughnut to coincide with the z -axis in R^3 , and introduce cylindrical coordinates r, ϕ, z . With the Ansatz $\varphi(\mathbf{x}) = \varphi(r, z) + \phi$ and $\theta(\mathbf{x}) = \theta(r, z)$ the ϕ -coordinate separates, and we obtain a two-dimensional equation for $U(r, z)$ and $V(r, z)$. That such a separation of variables is possible follows directly from the $SO(2) \times SO(2)$ symmetry of the unknot configuration. The ensuing equations are defined on the half-plane $r \geq 0, z \in (-\infty, \infty)$, which at each ϕ determines our cross-sectional plane. Notice that the unknot solution is even in z , which halves the CPU time in our simulation.

In Figs 1–3 we describe an unknot vortex which is a result of over 50 h of CPU time with each time step taking about 9 minutes on a 9,600-triangle mesh. The mesh has been selected so that it is more dense at the boundaries and near the core of the vortex. In our simulation we find impressive convergence, allowing us to increase the time step by up to eight orders of magnitude while keeping the relative change in energy intact. The Hopf invariant is also very stable, for the final configuration we have $Q_H = 0.999996$...

The trefoil vortex ($Q_H = \pm 3$) shown in Fig. 4 is a result of almost 200 h in CPU time on a 21^3 cubic finite-element lattice with tri-cubic Hermite basis functions. Consequently the number of nodes is about the same as in the case of unknot, but due to a lack of any obvious reflection symmetry each time iteration now takes about 20 minutes of CPU time. The yellow centre in Fig. 4 corresponds to the core $\mathbf{x}(\lambda)$ of our initial configuration, which

has been determined using the energy concept developed by Simon¹⁸. The points in Fig. 4 have been evaluated using the piecewise cubic polynomial approximation obtained from the finite-element algorithm. We have verified that the nonhomogeneity in Fig. 4 reflects the underlying lattice structure that we have used in our simulation. Indeed, a 21^3 lattice is obviously too rough to describe our trefoil solution adequately, but unfortunately we do not have access to a computer that would allow us to use an essentially larger lattice. Nevertheless, we have found definite numerical stability in the sense that the final configuration in Fig. 4 has remained essentially intact under a large number of iterations. We view this stability as a strong evidence that we indeed have convergence towards a trefoil vortex solution.

We expect that our results will have numerous important applications, and we are investigating two examples in detail. As the order parameter in nematic liquid crystal and ^3He superfluid involves a unit three vector, a theoretical investigation of vortices in these materials might reveal the existence of stable knot-like structures. In particular, such structures should be experimentally accessible. Similarly, an investigation of knots in the context of spontaneously broken Yang–Mills–Higgs theories is of particular interest, as stable knot-like vortices may have important physical implications for the cosmology of the early Universe. □

Received 29 October 1996; accepted 4 March 1997.

- Thomson, W. H. *Trans. R. Soc. Edinb.* **25**, 217–220 (1869).
- Tait, P. G. *On Knots I II, III in Scientific Papers* Vol. I 273–347 (Cambridge Univ. Press, 1898).
- Atiyah, M. *The Geometry And Physics of Knots* (Cambridge Univ. Press, 1990).
- Kauffman, L. H. *Knots And Physics* (World Scientific, Singapore, 1993).
- Green, M. B., Schwarz, J. H. & Witten, E. *Superstring Theory* Vols I, II (Cambridge Univ. Press, 1987).
- Witten, E. *Commun. Math. Phys.* **121**, 351–375 (1989).
- Bowick, M. J., Chander, L., Schiff, E. A. & Srivastava, A. M. *Science* **263**, 943–945 (1994).
- Bauerle, C., Bunkov, Yu. M., Fisher, S. N., Godfrin, H. & Pickett, G. R. *Nature* **382**, 332–334 (1996).
- Ruutu, V. M. H. *et al. Nature* **382**, 334–336 (1996).
- Summers, D. W. *Notices of AMS* **42**, 528–537 (1995).
- Rebbi, C. & Soliani, G. *Solitons and Particles* (World Scientific, Singapore, 1984).
- Faddeev, L. D. & Takhtajan, L. A. *Hamiltonian Methods in the Theory of Solitons* (Springer, Berlin, 1987).
- Makhankov, V. D., Rybakov, Y. P. & Sanyuk, V. I. *The Skyrme Model: Fundamentals, Methods, Applications* (Springer, Berlin, 1993).
- Faddeev, L. D. *Quantization of Solitons* (Princeton preprint IAS-75-QS70, Institute for Advanced Study, 1970).
- Faddeev, L. D. in *Relativity, Quanta and Cosmology* Vol. 1 (eds Pantaleo, M. & De Finis, F.) 185–192 (Johnson Reprint, 1979).
- Vakulenko, A. F. & Kapitanski, L. V. *Dokl. Akad. Nauk USSR* **248**, 810–814 (1979).
- Sewell, G. *Adv. Eng. Software* **17**, 105–112 (1993).
- Simon, J. K. J. *Knot Theory Raminif.* **3**, 299–320 (1994).

Acknowledgements. We thank J. Hietarinta, T. Kärki, A. P. Niemi, K. Palo, J. Pitkäranta, J. Rahola, D. Riska, R. Scharein, G. Sewell, J. Simon, O. Tirkkonen and J. Tulkki for discussions and suggestions; S. Virtanen for his help with visualization and simulating the initial Ansatz for the trefoil; M. Gröhn for helping us with visualization; and the Center for Scientific Computing in Espoo, Finland for providing us with an access to their line of Silicon Graphics Power Challenge computers. L.F. was supported by the Russian academy of Science and the Academy of Finland; A.J.J. was supported by the Göran Gustafsson Foundation for Science and Medicine, and by the NFR.

Correspondence should be addressed to A.J.N. (e-mail: niemi@teorfyf.uu.se).

Encapsulation of bilayer vesicles by self-assembly

Scott A. Walker, Michael T. Kennedy & Joseph A. Zasadzinski

Department of Chemical Engineering and Materials Research Laboratory, University of California, Santa Barbara, California 93106-5080, USA

Vesicles of lipid bilayers have been investigated as drug-delivery vehicles for almost 20 years^{1–8}. The vesicles' interior space is separated from the surrounding solution because small molecules have only limited permeability through the bilayer. Single-walled (unilamellar) vesicles are made by a variety of non-equilibrium techniques, including mechanical disruption of lamellar phases by sonication or extrusion through filters, or chemical disruption by detergent dialysis or solvent removal⁵. These techniques do not,

however, allow the encapsulation of a specific volume, nor can they be used to encapsulate other vesicles. Here we show that molecular-recognition processes mediated by lipophilic receptors and substrates (here the biotin-streptavidin complex)⁹ can be used to produce a multicompartimental aggregate of tethered vesicles encapsulated within a large bilayer vesicle. We call these encapsulated aggregates vesosomes. Encapsulation is achieved by unrolling bilayers from cochleate cylinders^{5,10-12} which are tethered to the aggregate by biotin-streptavidin coupling. These compartmentalized vesosomes could provide vehicles for multi-component or multifunctional drug delivery^{2-4,6}; in particular, the encapsulating membrane could significantly modify permeation properties, or could be used to enhance the biocompatibility of the system.

The process described here takes advantage of both lipid phase behaviour and specific recognition via biotin-streptavidin coupling to create supramolecular aggregates. Lipid molecules in aqueous solution first self-assemble into vesicles, then the vesicles are assembled into larger aggregates via specific ligand-receptor interactions⁹. The aggregates are encapsulated within an outer bilayer by first attaching them to cochleate cylinders of phosphatidylserine bilayers complexed with calcium^{5,10-13}. Cochleate cylinders are multilamellar lipid tubules formed spontaneously by certain negative phospholipids such as phosphatidylserine in the presence of Ca^{2+} (refs 10-13). Finally, the cylindrically wrapped phosphatidylserine bilayers are unrolled around the vesicle aggregate to form a closed outer membrane by lowering the calcium concentration (Fig. 1). In principle, almost any sub-micrometre to micrometre

sized object that can be labelled with the appropriate ligand can be encapsulated in a bilayer membrane via this process. Both the interior^{2,8} and exterior¹² bilayer compositions, the size and size distribution of the interior vesicles^{2,8,14}, the overall size of the vesosome, the nature of the vesicle tethers, and the type of additives to the outer bilayer (such as polymers for steric stabilization¹⁵ or specific recognition^{16,17}), can be optimized to create an extremely versatile structure.

To create a vesosome, large aggregates of unilamellar vesicles were prepared as described in ref. 9 using biotin-streptavidin tethers. But for applications in drug delivery through the circulation, compact, near-spherical aggregates $\leq 0.5 \mu\text{m}$ in diameter are preferred^{2,8,14}. While it is possible to inhibit the initial aggregation by controlling the biotin-streptavidin stoichiometry or by using charged vesicles¹⁸, the aggregate size distribution can be efficiently reduced by a second extrusion, this time through $1\text{-}\mu\text{m}$ pore size Nucleopore filters to produce a dispersion of compact 'sized' vesicle aggregates with diameters ranging from 0.3 to $1.0 \mu\text{m}$ (Fig. 2). The biotinylated lipid ('biotin-lipid') concentration was chosen to leave uncomplexed ligands on the vesicle surface. The sized vesicle aggregates were stable for weeks and experienced minimal reaggregation, indicating that no uncomplexed streptavidin was present¹⁸.

To encapsulate the vesicle aggregate, the aggregates were coupled to cochleate cylinders, again using the biotin-streptavidin interaction. The cochleate cylinders unroll to form large unilamellar vesicles (LUVs) on removal of Ca^{2+} by a chelating agent such as ethylenediaminetetraacetic acid (EDTA). 1,2-dioleoylphosphatidylserine (DOPS) cochleate cylinders, labelled with small amounts of biotin-lipid, were prepared as in ref. 10, then sufficient streptavidin was added to the cylinders to fully saturate all of the surface biotin-lipids. Sufficient streptavidin was added to the cylinders to fully saturate all of the surface biotin-lipids. Hence, the streptavidin-saturated cylinders can bind only to the vesicle aggregates when mixed together. In turn, by removing Ca^{2+} , the cylinder is unrolled while attached to the aggregate, encapsulating the aggregate within a DOPS bilayer. The cylinder engulfs the vesicle aggregate to form a closed membrane surface, facilitated by the formation of biotin-streptavidin bonds (Figs 1, 3).

Freeze-fracture transmission electron microscopy (FF-TEM) of samples prepared one day after mixing revealed that encapsulation of the vesicle aggregates had indeed taken place (Fig. 3). From Fig. 3a,

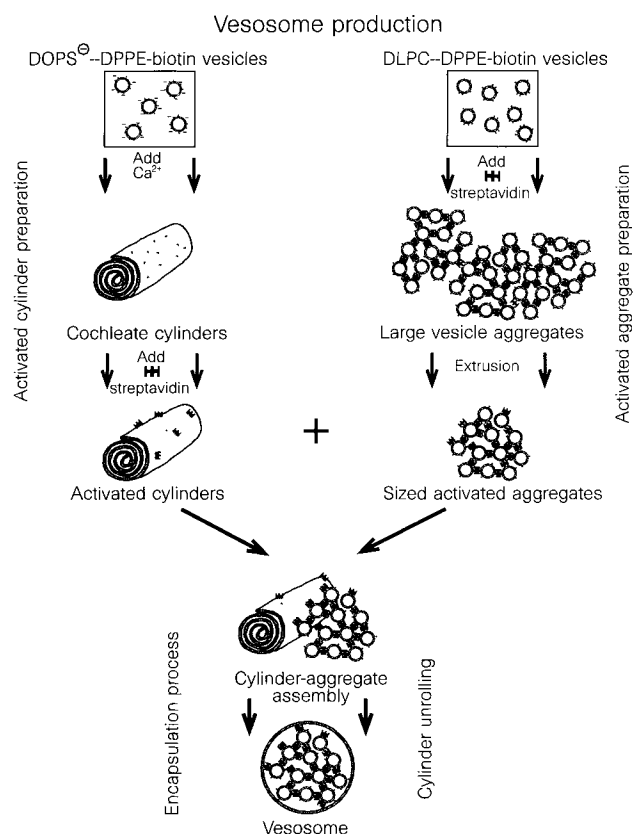


Figure 1 Vesosome production. Their construction involves two parallel processes, one being the production of vesicle aggregates, the other being the production of cochleate cylinders. Activated cylinders and sized aggregates are finally mixed. The cylinders and vesicle aggregates bind together, and the cylinder unrolls around the vesicles, attached by biotin-streptavidin linkages.

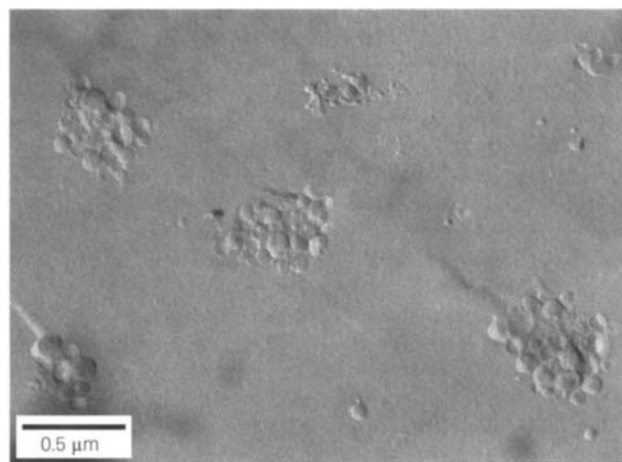


Figure 2 Freeze-fracture transmission electron micrograph of compact vesicle aggregates, $\sim 0.5 \mu\text{m}$ in diameter, produced after extruding large aggregates prepared as in ref. 9 through $1\text{-}\mu\text{m}$ Nucleopore filters. This produced a dispersion of compact 'sized' vesicle aggregates with diameters ranging from 0.3 to $1.0 \mu\text{m}$. Isolated monomer vesicles and small vesicle aggregates also occurred but could be removed by centrifugation or gel filtration.

it is clear that a single bilayer encapsulated the entire vesicle aggregate, consistent with the 'unrolling' of a single cochleate cylinder. Most of the structures visible in the freeze-fracture images were spheroidal structures, 0.5–5 μm in diameter, which were either large unilamellar vesicles or vesosomes that fractured around the periphery instead of through the centre (cross-fractured); it was difficult to distinguish between LUVs and vesosomes unless the structures cross-fractured. Some cylinders, unencapsulated vesicle aggregates, and free vesicles (100 nm) were also observed; we made no effort to isolate the vesosomes after the self-assembly steps.

To completely remove the Ca^{2+} , excess EDTA was added to the mixtures. The cloudy, opaque solution immediately turned greyish and more transparent, indicative of LUV suspensions. FF-TEM of samples taken five hours after EDTA addition revealed that each of these solutions contained many vesosomes (Fig. 3b) and LUVs (1–5 μm) (or vesosomes that had not cross-fractured). Also, there were several unencapsulated sized vesicle aggregates and monomer vesicles (left over from the extrusion step), but no cylinders. In Fig. 3b, the structure on the right is a cross-fractured vesosome; the structure on the left, which may be adhering to the vesosome, appears to have some type of protrusions on its surface. These protrusions are probably bulges in the bilayer from the encapsulated internal vesicles. The internal contents of the vesosomes resembled the sized vesicle aggregates (Fig. 2) in both size ($\sim 0.5 \mu\text{m}$) and aggregation state (dense, compact) and the vesicles in the interior aggregates were $\sim 100 \text{ nm}$ in diameter (Fig. 3).

Vesosomes were not present in any of the precursor solutions, nor were they seen in control experiments. The solutions of cylinders saturated with streptavidin (before and after EDTA addition) did not show any unusual characteristics due to the presence of streptavidin except that the cylinders seemed to disperse more readily, possibly due to the bound streptavidin sterically preventing the cylinders from aggregating¹⁹. The solutions of sized vesicle aggregates also did not exhibit any unusual structures. No LUVs were present in either of these solutions. The control experiments (separate additions of Ca^{2+} and EDTA to the sized vesicle aggregates) did not show structures resembling vesosomes.

Our images show that a step-wise process of individual self-

assembly steps consistent with the schematic shown in Fig. 1 produces a multicomponent, multicompartiment aggregate. The process of encapsulation is simple and probably generic to any small aggregate that can be labelled with the appropriate ligand. The combination of cochleate cylinders and ligand–receptor interactions is sufficient to produce a spontaneous encapsulation of the vesicle aggregates in a gentle way, thereby retaining the integrity of the interior membranes (Fig. 3).

We have not yet developed a good method quantifying the efficiency of entrapment. From our micrographs, it appears to be of order 5–15% efficient at present, without optimization of any of the steps. This level is clearly good enough to show the general technique works, but probably not good enough to do economical drug delivery. However, we are currently working on separation techniques (to remove uncomplexed vesicles) and more efficient binding techniques (to ensure one vesicle aggregate to one cochleate cylinder) which should get our efficiency to the 50% range.

There are many potential benefits of the vesosome over typical unilamellar vesicles used in applications. Important membrane functions can be divided among two or more vesicle membranes of different compositions and dimensions. The permeation rate, the membrane charge, specific recognition molecules, steric stabilizers, membrane rigidity and phase transition temperatures all play a role in the optimization of vesicle applications^{2–4,8,20}. With the vesosome structure, these often incompatible attributes can be divided among the various membranes. The cochleate cylinders that form the exterior membrane can be made from a variety of anionic lipids and mixtures of anionic and neutral lipids such as cholesterol^{5,12}, and the interior vesicles can have virtually any lipid composition or size. Many of the recent advances in vesicle-based drug delivery^{8,20} can be incorporated into the vesosome structure. For example, extended blood circulation times have been achieved by incorporating steric stabilizers, such as gangliosides, or polyethylene glycol coupled to a lipid (PEG–lipid) to the outer surface of the vesicles^{2,14,15,21}. Uster *et al.*¹⁵ have shown that equivalent steric stabilization can be achieved by adding PEG–lipids to pre-formed membranes, such as the outer membrane of the vesosome, by partitioning from an aqueous PEG–lipid micellar phase. The greatest benefit to this step-wise process of vesosome construction is the

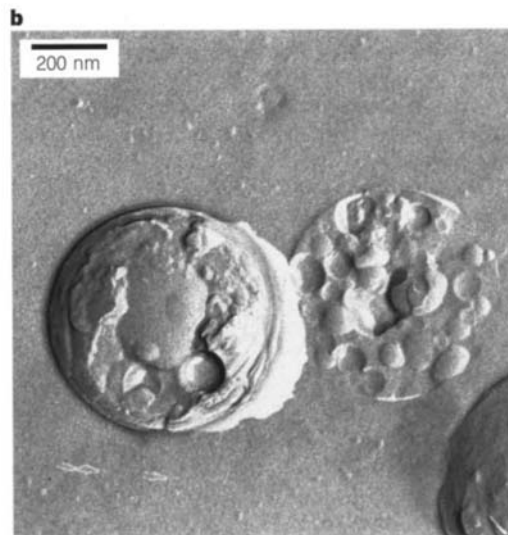
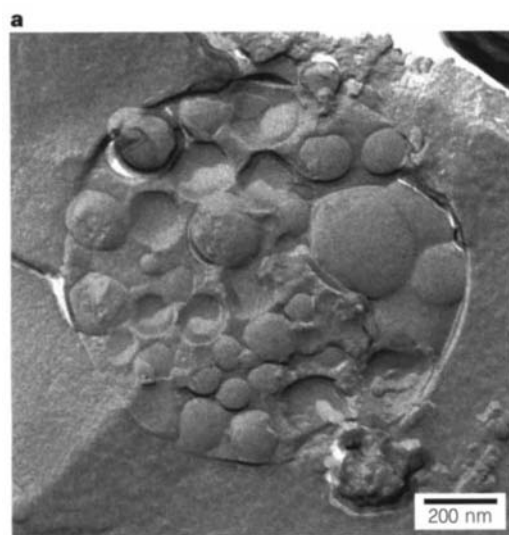


Figure 3 **a**, Freeze-fracture transmission electron micrograph of a typical vesosome formed after mixing streptavidin-coated cylinders and biotin-lipid coated aggregates (pre-EDTA; see text). There is one outer bilayer; interior vesicles appear to be specifically aggregated and $\sim 0.1 \mu\text{m}$ in diameter. **b**, Freeze-fracture transmission electron micrograph of adjacent vesosomes after addition

of EDTA to the solution shown in **a**. The vesosome on the left may have fractured around the periphery. The structures observed were little changed in comparison to the pre-EDTA samples. However, the single bilayer encapsulation of undamaged interior vesicles is obvious.

flexibility it allows in optimizing bilayer composition and aggregate size, at each step of the self-assembly process. □

Methods

Interior vesicle aggregates. Unilamellar vesicles were prepared by mixing dilauroylphosphatidylcholine (DLPC; Avanti Polar Lipids, Alabaster, Alabama) and dipalmitoylphosphatidylethanolamine-conjugated biotin (DPPE-biotin; Molecular Probes, Eugene, Oregon) at 0.16 mol% total lipid in chloroform, evaporating the solvent, then hydrating the lipid film with aqueous buffer (100 mM NaCl, 50 mM TES, and 0.02 wt% NaN₃ balanced to pH 7.2) for 48 h at 37 °C to form a 30 mg ml⁻¹ dispersion of multilamellar vesicles. The solution was put through five freeze-thaw (liquid nitrogen-50 °C water bath) cycles followed by ten high-pressure extrusions through two stacked 100-nm-pore polycarbonate Nuclepore filters (Corning Costar, Cambridge, Massachusetts) to produce a relatively monodisperse dispersion of 100-nm unilamellar vesicles (Fig. 1a). A 100-nm vesicle of this composition contains ~80 DPPE-biotins protruding from the monolayer.

To aggregate the vesicles, sufficient (0.63 mg ml⁻¹) streptavidin (Molecular Probes, Eugene, Oregon) in the same buffer was added to produce an overall biotin-streptavidin ratio of 15:1; however, the ratio of biotin on the outside of the vesicle available for binding to streptavidin was 8:1. As streptavidin has four distinct binding sites for biotin, the ratio of exposed biotins to binding sites was 2:1, meaning there are excess surface ligands (Fig. 1). The addition of streptavidin solution diluted the dispersion of unilamellar vesicles to 20 mg per ml total lipid. Within an hour, the 20 mg per ml ULV/streptavidin suspension changed from clear and bluish to opaque and cloudy-white, indicating that vesicle aggregates were forming⁹. The aggregates were filtered under pressure through 1.0-µm Nucleopore filters to produce the sized aggregates shown in Fig. 2.

Cochleate cylinders. Cochleate cylinders were prepared by first making a dispersion of 100-nm unilamellar vesicles containing 10 mg ml⁻¹ of 1,2-diolsphosphatidylserine (DOPS; Avanti Polar Lipids, Alabaster, Alabama) with 0.16 mol% DPPE-biotin as previously described⁹. Equal (1 ml) volumes of the DOPS/DPPE-biotin vesicle solution (10 mg ml⁻¹) and a 6 mM CaCl₂ (Sigma, St Louis, Missouri) buffer solution were mixed. Immediately after mixing, the solution turbidity increased, indicating that cochleate cylinders had formed. Freeze-fracture transmission electron microscopy (not shown) confirmed that cochleate cylinders had formed, indicating the added DPPE-biotin did not alter the cochleate structure. 35 µl of 0.63 mg ml⁻¹ streptavidin solution was injected into 1 ml of the cochleate cylinder solution and allowed to incubate for one day to fully saturate the biotin-lipids at the cylinder surface.

Vesosome assembly. The sized vesicle aggregates and the cylinders were mixed at a 1:1 DLPC:DOPS mole ratio: 1.0 ml of the 5 mg ml⁻¹ DOPS cylinders (6.2 µmol of DOPS) was added to 0.19 ml of the 20 mg ml⁻¹ DLPC vesicles (6.2 µmol of DLPC). To remove calcium ions, 0.44 ml of 5 mM EDTA was added to 0.5 ml of the mixture, resulting in a solution containing 4.2 mg ml⁻¹ DOPS lipid and 3.2 mg ml⁻¹ DLPC lipid. Freeze-fracture TEM samples were prepared by standard techniques²² after one day of incubation before adding EDTA (Fig. 3a), and 5 h of incubation after adding EDTA (Fig. 3b).

Received 23 December 1996; accepted 25 March 1997.

1. Bangham, A. D., Standish, M. M. & Watkins, J. C. Diffusion of univalent ions across lamellae of swollen phospholipids. *J. Mol. Biol.* **13**, 238–252 (1965).
2. Lasic, D. D. *Liposomes: from Physics to Applications* (Elsevier, Amsterdam, 1993).
3. Gregoriadis, G. *Liposomes as Drug Carriers—Recent Trends and Progress* (Wiley, New York, 1988).
4. Fendler, J. *Membrane Mimetic Chemistry* (Wiley, New York, 1983).
5. Szoka, F. & Papahadjopoulos, D. Comparative properties and methods of preparation of lipid vesicles (liposomes). *Annu. Rev. Biophys. Biochem.* **9**, 467–508 (1980).
6. New, R. R. C. (ed.) *Liposomes: a Practical Approach* (Oxford Univ. Press, 1990).
7. Spector, M. S., Zasadzinski, J. A. & Sankaram, M. B. Topology of multivesicular liposomes, a model biliquid foam. *Langmuir* **12**, 4704–4708 (1996).
8. Allen, T. M., Hansen, C. B. & Lopes de Menezes, D. E. Pharmacokinetics of long-circulating liposomes. *Adv. Drug Delivery Rev.* **16**, 267–284 (1995).
9. Chiruvolu, S., Walker, S., Leckband, D., Israelachvili, J. & Zasadzinski, J. Higher order self-assembly of vesicles by ligand-receptor interactions. *Science* **264**, 1753–1756 (1994).
10. Papahadjopoulos, D., Vail, W. J., Jacobson, K. & Poste, G. Cochleate lipid cylinders formation by fusion of unilamellar lipid vesicles. *Biochim. Biophys. Acta* **394**, 483–491 (1975).
11. Papahadjopoulos, D., Vail, W. J., Pangborn, W. A. & Poste, G. Studies on membrane fusion II: induction of fusion in pure phospholipid membranes by calcium ions and other divalent ions. *Biochim. Biophys. Acta* **448**, 265–283 (1976).
12. Coosens, J. R. & Rand, R. P. Structural effects of neutral lipids on divalent cation-induced interactions of phosphatidylserine-containing bilayers. *Biophys. J.* **68**, 1009–1018 (1995).
13. Papahadjopoulos, D. et al. Studies on membrane fusion III: the role of calcium-induced phase changes. *Biochim. Biophys. Acta* **465**, 579–598 (1977).

14. Allen, T. M. & Chonn, A. Large unilamellar liposomes with low uptake into the reticuloendothelial system. *FEBS Lett.* **223**, 42–46 (1987).
15. Uster, P. S. et al. Insertion of poly(ethylene glycol) derivatized phospholipid into pre-formed liposomes results in prolonged in vivo circulation time. *FEBS Lett.* **386**, 243–246 (1996).
16. Zalipsky, S., Hansen, C. B., Lopes de Menezes, D. E. & Allen, T. M. Long-circulating, polyethylene glycol grafted immunopolisomes. *J. Controlled Release* **39**, 153–161 (1996).
17. Wong, J. Y., Kuhl, T. L., Israelachvili, J. N., Mullah, N. & Zalipsky, S. Direct measurement of a tethered ligand-receptor interaction potential. *Science* **275**, 820–822 (1997).
18. Walker, S. A. thesis, Univ. California, Santa Barbara, (1996).
19. Kennedy, M. T. thesis, Univ. California, Santa Barbara (1998).
20. Tardi, P. G., Boman, N. L. & Cullis, P. R. Liposomal doxorubicin. *J. Drug Targeting* **4**, 129–140 (1996).
21. Papahadjopoulos, D. et al. Sterically stabilized liposomes: improvements in pharmacokinetics and antitumour therapeutic efficacy. *Proc. Natl Acad. Sci. USA* **88**, 11460–11464 (1991).
22. Chiruvolu, S., Naranjo, E. & Zasadzinski, J. A. in *Structure and Flow in Surfactant Solutions* Ch. 5 (eds Herb, C. A. & Prud'homme, R. K.) (Am. Chem. Soc., Washington DC, 1994).

Acknowledgements. We thank J. N. Israelachvili for suggestions on the vesicle inside a vesicle concept and the name vesosome. This work was supported by the NIH, the NSF, and the Materials Science and Engineering Research Center program of the NSF.

Correspondence and requests for materials should be addressed to J.A.Z. (e-mail: gorilla@engineering.ucsb.edu).

Temperature effects on the acidity of remote alpine lakes

Sabine Sommaruga-Wögrath*†, Karin A. Koinig*, Roland Schmidt†, Ruben Sommaruga*, Richard Tessadri§ & Roland Psenner*

* Institute of Zoology and Limnology, University of Innsbruck, 6020 Innsbruck, Austria

† Institute of Limnology, Austrian Academy of Sciences, 5310 Mondsee, Austria

§ Institute of Mineralogy and Petrography, University of Innsbruck, 6020 Innsbruck, Austria

Climate variations and changes in sulphur and nitrogen deposition from the atmosphere influence the acid-base balance of sensitive lakes in a complex and site-specific way^{1–3}. For example, although lakes in several regions have shown a decline in sulphate concentration following reductions in atmospheric sulphate deposition^{4–6}, the expected recovery of pH and alkalinity has not always taken place, implicating an additional response to changes in the local climate. Here we report a study of 57 remote alpine lakes which shows that, between 1985 and 1995, lake pH and the concentration of sulphate, base cations and silica have increased, whereas inorganic nitrogen concentrations have decreased. This contrasts with atmospheric input trends, which have led to a decrease in sulphate and a slight increase in nitrogen deposition over the same period^{7,8}. We propose that the changes in lake chemistry are therefore likely to be caused by enhanced weathering and increased biological activity resulting from an increase in air temperature of about 1 °C since 1985. Our analysis of an alpine lake core covering a 200-year period provides further evidence for a strong positive correlation between pH and mean air temperatures, and thus for the high sensitivity of lakes at high altitudes and high latitudes to climate warming. In these remote locations, temperature effects, rather than acid deposition, appear to dominate changes in lake acidity.

We studied 57 low-alkalinity high-mountain lakes in glaciated and non-glaciated catchments, situated between 2,000 and 2,900 m above sea level (m.a.s.l.) on the northern (North Tyrol) and southern slope (East Tyrol, Carinthia) of the eastern Alps. The area is characterized by granites and gneisses of high sensitivity to acid deposition⁹. Soils are poorly developed with sparse vegetation, especially at very high altitudes where large portions (70–90%) of the catchments consist of bare rock. Samples were collected during the autumn overturn in 1985 and 1995 and analysed for pH,

† Present address: Institute of Hygiene, University of Innsbruck, 6010 Innsbruck, Austria.

conductivity, alkalinity, silica, dissolved organic carbon (DOC), dissolved organic nitrogen (DON), major anions and cations, and phosphorus. In Schwarzsee ob Sölden (SOS, 2,796 m.a.s.l.), a sediment core, covering the period from 1778 to 1991, was collected for analyses of fossil diatoms and chrysophytes, fossil pollen, and geochemistry. Monitoring¹⁰ of SOS began in 1984 and is still going on.

Since the minimum in 1885, air temperatures have warmed by $\sim 2^\circ\text{C}$ at mountain stations in the central Alps^{11,12}; 1°C of this change has occurred since 1985. The overall increase is more than twice that reported for the Northern Hemisphere average (Fig. 1a; compare also Fig. 4 for running averages) but comparable to the mean calculated from 20 stations in Austria, located predominantly at lower altitudes¹³. As in the Swiss Alps¹¹, total annual precipitation in the central Alps has not changed significantly since 1980 (Fig. 1b). Sulphate deposition has decreased since 1985, but inorganic nitrogen (dissolved inorganic nitrogen: DIN = nitrate + nitrite + ammonium nitrogen) increased slightly during this period (Fig. 1c).

In contrast to deposition trends, base cations and SO_4^{2-} increased between 1985 and 1995 in most lakes, while inorganic nitrogen (DIN) decreased (Fig. 2). DON concentrations increased in all but four lakes. The pH (Fig. 2a) increased by an average of 0.25 units, calculated from differences in H^+ concentrations. Alkalinity (Fig. 2b) increased in 63%, but decreased in 28%, of the lakes, with an average increase of

$\sim 10 \mu\text{equiv. l}^{-1}$. The mean increase in silica of $10 \pm 20 \mu\text{mol l}^{-1}$ between 1985 and 1995 compares fairly well with an annual increase of $1.4 \mu\text{mol l}^{-1}$ in the SOS (Fig. 3) from 1984 to the end of 1996. The SOS has undergone a gradual increase in sulphate ($4.2 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$) and base ion concentrations ($4.4 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$), whereas nitrogen decreased by $0.4 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$ and chloride showed no significant change.

One of the most obvious effects of higher mean air temperatures on high mountain lakes is a shorter period of snow and ice cover, which improves exchange of gases and nutrients, wind-driven circulation and light conditions^{14,15}. Around 1900, for example, and during some years in the cooler period around 1970, the slopes in the catchment area of the SOS were permanently covered by snow. In 1995 these slopes were snow-free in summer, and the resultant increase in temperature exceeded by far the average difference of 1 or 2°C in air temperature (Fig. 1a), especially on south-facing slopes. Since 1980, glaciers in the eastern Alps have retreated continuously and 90% of the glaciers examined ($n = 96$) in 1995 have suffered a substantial loss in area¹⁶. This has created relatively large surfaces covered with fine glacial flour which provide, in combination with higher temperatures and liquid water, a reactive environment for intense chemical weathering and mineral dissolution¹⁷ (Fig. 2c).

In the last 10 years, SO_4^{2-} concentration has increased in 80% of

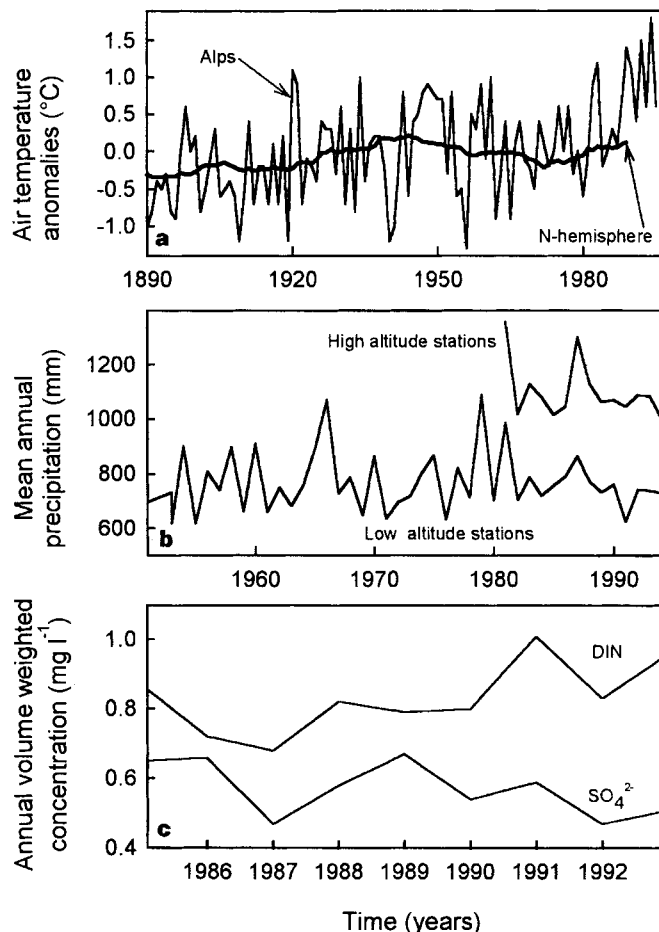


Figure 1 Temperature, precipitation and deposition trends in Alpine stations. **a**, Annual mean air temperature anomalies from five stations of the eastern Alps, and the Northern Hemisphere average; calibration period, 1901–50. **b**, Mean annual precipitation from five mountain stations. **c**, Trends in annual volume weighted sulphate and DIN (dissolved inorganic nitrogen: nitrate, nitrite and ammonium) concentrations in precipitation.

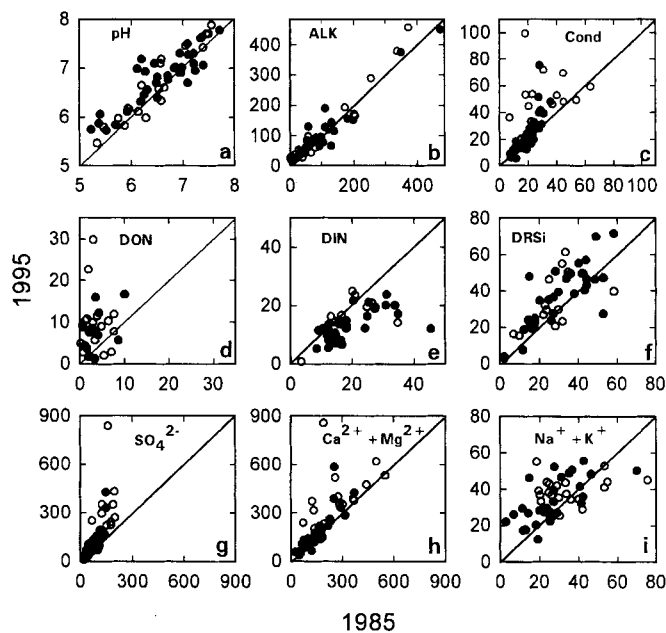


Figure 2 Comparison of chemical parameters from 57 high mountain lakes measured in 1985 and 1995. Open and closed circles represent lakes with glaciated and non-glaciated catchments, respectively. The solid lines represent the 1:1 ratio. **a**, pH; **b**, alkalinity ($\mu\text{equiv. l}^{-1}$); **c**, conductivity ($\mu\text{S cm}^{-1}$); **d**, dissolved organic nitrogen, DON (total dissolved nitrogen – DIN, $\mu\text{mol l}^{-1}$); **e**, dissolved inorganic nitrogen, DIN (nitrate + nitrite + ammonium, $\mu\text{equiv. l}^{-1}$); **f**, dissolved reactive silica, DRSi ($\mu\text{mol l}^{-1}$); **g**, sulphate ($\mu\text{equiv. l}^{-1}$); **h**, calcium + magnesium ($\mu\text{equiv. l}^{-1}$); **i**, sodium + potassium ($\mu\text{equiv. l}^{-1}$).

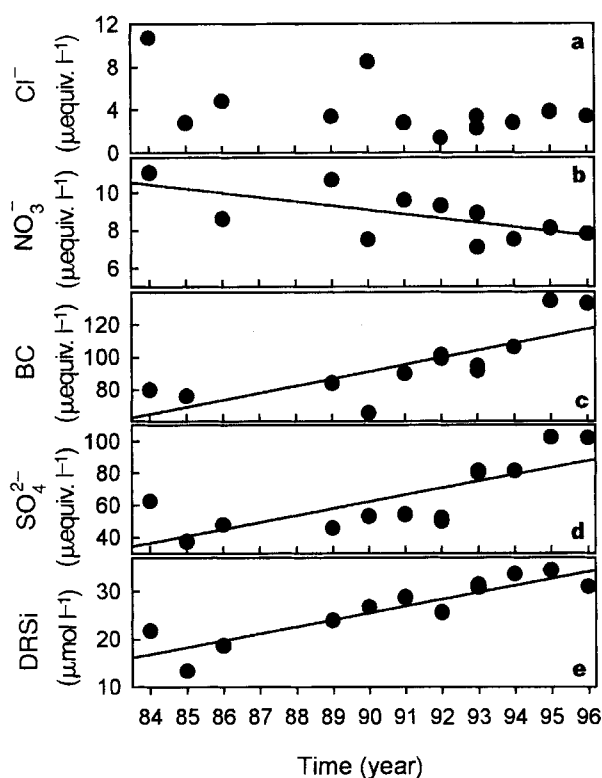


Figure 3 Changes in ion and dissolved reactive silica concentration in Schwarzsee ob Sölden under open lake conditions (summer) from 1984 to the end of 1996, and annual trends calculated from the slope of linear regressions. **a**, chloride. **b**, Nitrate, $-0.4 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$ ($r^2 = 0.48$, $P < 0.01$). **c**, Base cations (BC), $4.4 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$ ($r^2 = 0.60$, $P < 0.01$). **d**, Sulphate, $4.2 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$ ($r^2 = 0.56$, $P < 0.01$). **e**, Dissolved reactive silica (DRSi), $1.4 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$ ($r^2 = 0.79$, $P < 0.0001$).

the investigated lakes; in lakes surrounded by glaciers, concentrations were up to 400% higher than in 1985 (Fig. 2g). Pulses of sulphate-rich water low in alkalinity following drought or due to desorption from mineral surfaces have been documented at several acidified sites in North America^{18–21}. However, as annual precipitation was fairly constant in the eastern Alps during the last decade (Fig. 1b) and lakewater alkalinity increased slightly (Fig. 2b), we assume that dissolution of easily soluble calcium and magnesium sulphates in the catchment is the main factor responsible for the SO_4^{2-} increase. This is supported by the fact that the increase in SO_4^{2-} was generally well balanced by Ca^{2+} and Mg^{2+} both in non-glaciated ($r^2 = 0.81$, $P < 0.05$) and glaciated catchments ($r^2 = 0.97$, $P < 0.05$). Pyrite oxidation, which produces sulphuric acid and thus enhances the weathering of silicate or carbonate minerals, could also explain increased concentrations of SO_4^{2-} and base cations, but the comparably small increase in silica and alkalinity (Fig. 2f, b) suggests that most alkaline-earth and SO_4^{2-} ions are dissolved concomitantly (Fig. 2h). Monitoring data from the SOS (Fig. 3) showed similar rates of base-cation and sulphate increase (4.4 and $4.2 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$, respectively), but a much lower rate for silica ($1.4 \mu\text{equiv. l}^{-1} \text{ yr}^{-1}$) increase. The increase in silica, paralleled by equimolar amounts of sodium and potassium (Fig. 2f, i), indicates enhanced silicate weathering in ~80% of the lakes. Studies in granite and gneiss catchments have shown that weathering rates may intensify by an order of magnitude when temperature is increased from 0 to 25 °C while precipitation is kept constant²². Such temperature increases are possible in summer on sunny slopes which were previously snow-covered throughout the year.

Contrary to precipitation trends, DIN concentrations declined in 81% of the lakes. A shorter period of lake ice-cover, like that

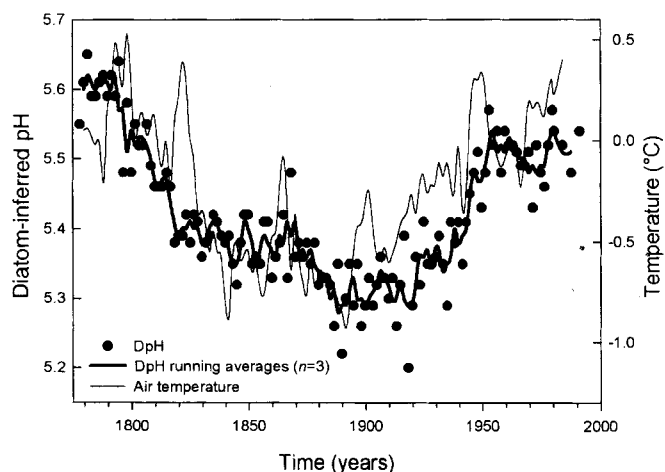


Figure 4 Diatom-inferred lakewater pH (DpH) from the sediment core of Schwarzsee ob Sölden and mean air temperature in Austria for the period 1778–1991. Air temperatures are the running averages ($n = 9$) of annual means from 20 stations¹³.

observed in the SOS, probably improved the conditions for phytoplankton growth by increasing light availability, temperature and nutrient input. The enhanced biological production would lead to faster recycling of nitrogen, which explains the lower DIN, and higher DON, concentrations found in almost all lakes (Fig. 2d, e). This assumption is supported by the relatively large increase in pH (0.25 units, compared to an alkalinity increase of only $10 \mu\text{equiv. l}^{-1}$) indicating higher productivity. The alternative explanation, that the above conditions promoted higher productivity of terrestrial plants, is less probable as the soils are poorly developed and low in organic content.

Further evidence for the dominant role of temperature on pH comes from palaeolimnological investigations in the SOS. In this lake, the diatom-inferred pH (DpH) has gradually increased since 1900 by 0.35 units (Fig. 4), following closely the mean air temperature of Austria. Statistical analyses of the mean air temperature and of the stratigraphical pollen, carbon, nitrogen and DpH data show a significant correlation between temperature and DpH, with 48% of the DpH variation explained by temperature and an additional 13% by carbon and nitrogen concentrations. A similar correlation was found for two lakes of the southern slope of the Alps, where acid precipitation has led to a decoupling of the pH–temperature relationship observed before 1900 (ref. 23). In the SOS, pH trends can be approximated by the formula $\text{pH} = 5.46 + 0.19 \text{ Temperature (°C)}$ ($r = 0.68$, $P < 0.001$, $n = 128$). Notwithstanding the considerable differences in altitude and bedrock geology of the 57 lakes studied, the mean pH increase of 0.25 units between 1985 and 1995 corresponds surprisingly well with the predictions obtained from palaeolimnological data.

Our results show that climate warming is a determinant factor for the water chemistry of remote alpine lakes, and that it can mitigate or even compensate for the anticipated effects of acid deposition. A compensating effect on the expected recovery of pH and alkalinity following reduced sulphur deposition has previously been attributed to an increase in nitrogen loads²⁴, a decrease in base-cation deposition^{25–28} and reoxidation of reduced sulphur in lake sediments² or runoff from catchments^{18,19}. Changes in lakewater chemistry due to climate warming have been suggested before^{14,20}. But while these earlier studies identified climatic effects due to drought-induced changes in water retention times¹⁴ and groundwater flow²⁰, our data indicate that enhanced weathering and biological activity are the main mechanisms that cause the trends in ionic composition, nitrogen and silica concentration, and pH observed in the study area. This finding implies that aquatic ecosystems at both high altitudes and

high latitudes might respond with extreme sensitivity to climate warming. In addition, our results further illustrate that the quantity and direction of change due to climate warming depend on site-specific processes and are, for this reason, not readily predictable for aquatic systems in different climatic and geographical areas. □

Methods

Water analysis. Water samples were taken, during autumn overturn, from several depths with a Patalas-Schindler sampler. Analytical methods: alkalinity, Gran titration; major anions and cations, ion chromatography; total dissolved nitrogen, oxidation by potassium peroxodisulphate and reduction on a Cd column; dissolved reactive silica, oxalic acid/molybdate method. Internal quality controls of the data were made by ionic charge balance and by comparing the measured and calculated conductivity. External quality assurance was provided by regular participation in international ring tests.

Palaeolimnology. Recent sediments were obtained with a modified Kajak sampler (outer diameter = 6 cm) and cut in 0.25 cm slices. Fossil diatoms were extracted using standard techniques²⁹. DpH was calculated using weighted average regression and a regional alpine data set³⁰. The average annual sedimentation rate of the core, dated by ¹³⁷Cs and ²¹⁰Pb, was 1.5 mm ± 0.1 mm over the last 100 years. For sediment layers older than 100 years the dating error may be substantial.

Received 14 August 1996; accepted 4 March 1997.

- Schindler, D. W., Curtis, P. J., Parker, B. R. & Stainton, M. P. Consequences of climate warming and lake acidification for UV-B penetration in North American boreal lakes. *Nature* **379**, 705–708 (1996).
- Yan, N. D., Keller, W., Scully, N. M., Lean, D. R. S. & Dillon, P. J. Increased UV-B penetration in a lake owing to drought-induced acidification. *Nature* **381**, 141–143 (1996).
- Wright, R. F. & Schindler, D. W. Interaction of acid rain and global changes: effects on terrestrial and aquatic ecosystems. *Wat. Air Soil Pollut.* **85**, 89–99 (1995).
- Gunn, J. M. & Keller, W. Biological recovery of an acid lake after reductions in industrial emissions of sulphur. *Nature* **345**, 431–433 (1990).
- Ek, A., Grahn, O., Hultberg, H. & Renberg, I. Recovery from acidification in lake Örvattnet, Sweden. *Wat. Air Soil Pollut.* **85**, 1795–1800 (1995).
- Kros, J., Groeneweg, J. E., De Vries, W. & Van der Salm, C. Uncertainties in long-term predictions of forest soil acidification due to neglecting seasonal variability. *Wat. Air Soil Pollut.* **79**, 353–375 (1995).
- Puxbaum, H., Kovar, A. & Kaliua, M. in *Seasonal Snow Packs* (ed. Davies, T. D.) 273–297 (NATO ASI Ser. G28, Springer, 1991).
- Mosello, R., Tartari, G. & Tartari, G. A. Chemistry of bulk deposition at Pallanza (Northern Italy) during the decade 1975–1984. *Mem. Ist. Ital. Idrobiol.* **43**, 311–332 (1985).
- Psenner, R. Chemistry of high mountain lakes in siliceous catchments of the central Eastern Alps. *Aquat. Sci.* **51**, 108–127 (1989).
- Wögrath, S. & Psenner, R. Seasonal, annual and long-term variability in the water chemistry of a remote high mountain lake: acid rain versus natural changes. *Wat. Air Soil Pollut.* **85**, 359–364 (1995).
- Beniston, M., Rebetez, M., Giorgi, F. & Marinucci, M. R. An analysis of regional climate change in Switzerland. *Theor. Appl. Climatol.* **49**, 135–159 (1994).
- Nicolussi, K. & Bortenschlager, S. CO₂-Anstieg und Jahrringbreitrends bei *Pinus cembra* in den westlichen Zentralalpen Tirols (Österreich). *Ber. nat.-med. Verein Innsbruck* **82**, 61–77 (1995).
- Report on Anthropogenic Climate Changes and Possible Effects on Austria* (Austrian Academy of Sciences, Vienna, 1991).
- Schindler, D. W. *et al.* The effects of climatic warming on the properties of boreal lakes and streams at the Experimental Lakes Area, northwestern Ontario. *Limnol. Oceanogr.* **41**, 1004–1017 (1996).
- Schindler, D. W. *et al.* Effects of climatic warming on lakes of the central boreal forest. *Science* **250**, 967–970 (1990).
- Patzelt, G. Gletscherbericht 1993–94. *Mitt. Öster. Alpenvereins* **50**, 18–23 (1995).
- Chilrud, S. N., Pedrozo, F. L., Temporetti, P. F., Planas, H. F. & Froelich, P. N. Chemical weathering of phosphate and germanium in glacial meltwater streams: effects of subglacial pyrite oxidation. *Limnol. Oceanogr.* **39**, 1130–1140 (1994).
- Bayley, S. E., Schindler, D. W., Parker, B. R., Stainton, M. P. & Beatty, K. G. Retention and release of S from freshwater wetland. *Biogeochemistry* **17**, 191–204 (1992).
- Lazerte, B. The impact of drought and acidification on the chemical exports from a minerotrophic conifer swamp. *Biogeochemistry* **18**, 153–175 (1993).
- Webster, K. E., Newell, A. D., Baker, L. A. & Brezonik, P. L. Climatically induced rapid acidification of a softwater seepage lake. *Nature* **347**, 374–376 (1990).
- Webster, K. E. & Brezonik, P. L. Climate confounds detection of chemical trends related to acid deposition in upper midwest lakes in the USA. *Wat. Air Soil Pollut.* **85**, 1575–1580 (1995).
- White, A. F. & Blum, A. E. Effects of climate on chemical weathering in watersheds. *Geochim. Cosmochim. Acta* **59**, 1729–1747 (1995).
- Psenner, R. & Schmidt, R. Climate-driven pH control of remote alpine lakes and effects of acid deposition. *Nature* **356**, 781–783 (1992).
- Clair, T. A. *et al.* Regional precipitation and surface-water chemistry trends in southeastern Canada. *Can. J. Fish. Aquat. Sci.* **52**, 197–212 (1995).
- Camarero, L. *et al.* Acidification in high mountain lakes in central, southwest and southeast Europe (Alps, Pyrenees, Pirin). *Limnologia* **25**, 141–156 (1995).
- Miller, E. K., Blum, J. D. & Friedland, A. J. Determination of soil exchangeable-cation loss and weathering rates using Sr isotopes. *Nature* **362**, 438–441 (1993).
- Hedin, L. O. *et al.* Steep declines in atmospheric base cations in regions of Europe and North America. *Nature* **367**, 351–354 (1994).
- Likens, G. E., Driscoll, C. T. & Buso, D. C. Long-term effects of acid rain: response and recovery of a forest ecosystem. *Science* **272**, 244–246 (1996).
- Battarbee, R. W. in *Handbook of Holocene Paleoclimatology and Paleohydrology* (ed. Berglund, B. E.) 527–531 (Wiley, New York, 1986).
- Marchetto, A. & Schmidt, R. A regional calibration data set to infer lakewater pH from sediment diatom assemblages in alpine lakes. *Mem. Ist. Ital. Idrobiol.* **51**, 115–125 (1993).

Acknowledgements. We thank D. W. Schindler, R. Mosello, K. E. Webster, J. Birks, A. Marchetto, W. Wieser and C. Sturmbauer for improvements to the manuscript, and J. Franzoi and W. Müller for technical assistance in the field and in the laboratory. J. Wieser improved our English. This work was supported by the Ministry of Agriculture, the Austrian Science Foundation and the EU Environmental Programme ALPE 2.

Correspondence and request for materials should be addressed to R.P. (e-mail: roland.psenner@uibk.ac.at).

Structural features in a brittle-ductile wax model of continental extension

James N. Brune* & Michael A. Ellis†

* Seismological Laboratory, University of Nevada, Reno, Nevada 89557, USA

† Center for Earthquake Research and Information, University of Memphis, Memphis, Tennessee 38152, USA

Structural features produced during the rifting of continents depend on the layered rheological properties of the crust and lithosphere and, in particular, on the presence of any transitions between brittle and ductile behaviour¹. Here we use a wax model to explore the gross structural response of continental lithosphere under pure shear extension in the presence of a continuous brittle–ductile transition. The wax models were deformed under various boundary conditions to reflect a variety of different regions, most notably the Basin and Range province of North America. Our experiments show the development of listric normal faults, structures common to regions of continental extension. We also observe the formation of distributed and discrete rifting, and intrusion and occlusion of the upper brittle layer by the ductile lower layer. The factor controlling deformation style in each case appears to be the relative thickness of the brittle and ductile layers, although a relatively high rate of strain generally promotes discrete rifting.

Rheological layering in wax, including a continuous transition between brittle and ductile behaviour, may be made by generating a vertical temperature gradient giving a transition from brittle to ductile behaviour at some depth. Previous modelling of plate-boundary processes using wax involved a solid–fluid transition (rather than a brittle–ductile transition) and were able to demonstrate successfully the importance of the brittle layer in the production of ridge-transform structures at sites of ocean-floor spreading^{2,3}.

The first set of experiments was designed to test the influence of the relative thickness of brittle to ductile layers. A smoothly varying vertical thermal gradient was generated in the wax by heating from below (Fig. 1a), resulting in a relatively broad transition from a ductile to a brittle rheology (distinguishable by the relative opacity of the wax). The relative thickness of the brittle and ductile layers and the gradient of the transition was controlled by the time over which the upper surface was cooled (Fig. 1a). The upper portion of the slab remained brittle under all strain rates. Confining pressure is absent in these experiments and failure occurs by cracking rather than by faulting. If we adopt a coordinate system with x along the length of the slab, y perpendicular to x and within the slab, and z vertical, then the strain $e_{xx} > 0$ was imposed and both $\sigma_{yy} = \sigma_{zz} = 0$ and all shear stresses, σ_{ij} ($i \neq j$), were equal to zero on model surfaces. Of course, complex local variations in e_{xx} , e_{yy} , and e_{zz} resulted from the brittle–ductile transition. These boundary conditions are similar to those adopted in various numerical models^{4–6} and are likely to be similar to those in actively extending continents. In particular, if the double-layered model is considered analogous to a continental crust, then the infinitely stiff base under the ductile layer represents the strong upper mantle beneath the Moho. Alternatively, the slab

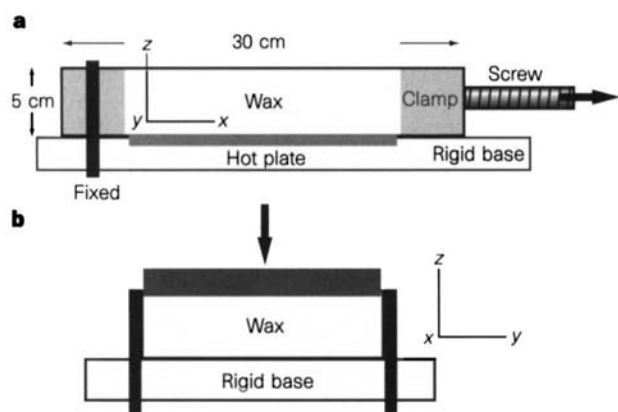


Figure 1 Schematics of experimental apparatus. **a**, A machined wax slab of hard microcrystalline paraffin wax is placed over a 24-cm-wide hot steel plate, and the upper surface is allowed to cool to room temperature. This generated a smoothly varying thermal (and therefore strength) vertical gradient. A sharper transition was achieved by bringing the entire slab to a high homologous temperature, thus making it wholly ductile, and placing an ice-pack on top of the slab. A thin (~0.1 mm) veneer of melt (~90°C) at the base of the model prevents the generation of significant shear stress across the lower boundary. No attempt was made to determine quantitatively the viscosity profile, because viscosity of the wax is highly temperature dependent. Mancktelow²⁹ has demonstrated, however, that in experimental conditions similar to ours, paraffin wax exhibits power-law creep behaviour analogous to that expected in the Earth's crust⁶. By trial and error the ductility of the lower layer was adjusted to be just ductile (not brittle) at the strain rates imposed, but still strong enough to rupture the upper brittle layer. The upper brittle layer was in all cases so cold that it could not deform in a ductile fashion. Coordinate axes are shown for reference to text. Experiments lasted ~1 h. Strain rates were of the order 10^{-4} s^{-1} , which means that 1 h of model time is ~1 Myr given a geological strain rate of 10^{-14} s^{-1} . **b**, A machined wax slab 16 cm in length, 8 cm wide and 4 cm thick, is compressed in pure shear between a confined trough. The slab is heated to a relatively high homologous temperature (~75–85°C) so that it is wholly ductile and then cooled from the top surface by an ice-pack. Experiments lasted ~1 min. Strain rates were of the order of 10^{-2} s^{-1} , so that 1 min is $\sim 2 \times 10^5 \text{ yr}$.

may represent the entire continental lithosphere subject to a high heat flow (eradicating any strong upper-mantle lid) or a more uniform oceanic lithosphere over a hotspot or incipient rift zone.

Multiple distributed rifting occurred when the ductile layer formed a significantly larger portion of the total thickness, and concentrated, discrete rifting occurred as the relative thickness of the brittle layer approached ~1/4 of the total thickness (Fig. 2). We explain this in terms of strain softening at the length scale of the double-layered slab, or in a geological context, at the length scale of the crust or entire lithosphere. That is, when the brittle portion is relatively thin, it supports an insignificant proportion of the total stress across a vertical plane. As the thin brittle layer fails, the force required to continue extension is reduced only slightly and layer-scale strain softening does not occur. By contrast, failure of a relatively thick brittle layer will significantly reduce the strength of the whole column, rapidly lowering the force required for continued extension, leading to model-scale strain softening and discrete or narrow rifting.

These basic results suggest that under similar rates of strain, relatively narrow rifts will form in conditions that favour a large fraction of brittle to ductile crust. That is, old, thick and cold crust might be expected to yield narrow rifts, whereas younger, thinner and warmer crust should give wider extensional zones. This is consistent with numerous observations in the East African, northern Red Sea, and Baikal rifts (narrow rifts) and in the Aegean Sea

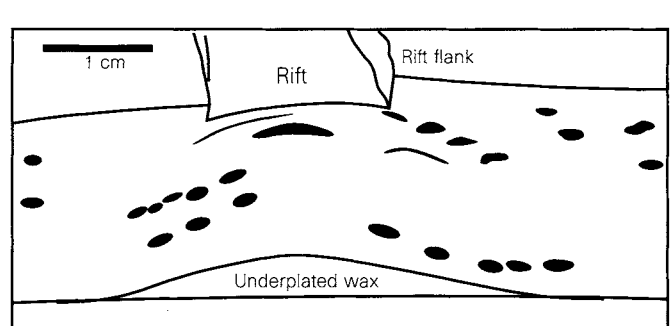


Figure 2 Photograph and line drawing of isolated rifting that occurred when the proportion of brittle to ductile layer thickness was large, so that failure of the upper brittle layer leads to slab-scale strain softening. The distribution of distorted air bubbles shows the strain to be concentrated in a narrow zone vertically through the layer. Highly attenuated bubbles show the high strain gradient between the upper and lower layers in the rift zone. During extension, molten wax collected under the uplifted rift zone, effectively underplating the rift as it later cooled. This process is considered to occur under oceanic³⁰ and continental rifts³¹.

and Neogene-aged structures in Basin and Range regions (wide rifts) (see ref. 7 and references therein).

A second set of experiments was designed to test the effect of confining pressure and larger finite strains (Fig. 1b). These were performed by confining a slab of wax in a rectangular metal trough and imposing a vertical compressional strain (using a metal plate) so that the wax was extruded perpendicular to the imposed strain (Fig. 1b). All surfaces were lubricated by grease to avoid the development of boundary shear stresses. The effect of confining the wax and compressing it in the trough created the approximate kinematic conditions $e_{xx} > 0$, $e_{yy} = 0$ and $e_{zz} < 0$, with expected complications due to edge effects and the brittle–ductile transition. The relative thickness of the brittle layer was controlled as in the first set of experiments.

The advantage of generating confining pressure is that the wax failed by faulting rather than by cracking. Brittle layers varied from a small fraction (~0.025) to a relatively large fraction (~0.25) of the whole slab before extension. Extension was accommodated by one or more listric normal faults (Fig. 3), each marked by slickensides. Listric normal faults are well documented in regions of continental extension (see, for examples, refs 8, 9) but their origin and existence remains a debated issue (for example, refs 10, 11). As far as we know, this is the first physical experiment that has generated continuous listric normal faults.

Various theoretical approaches to the development of low-angle

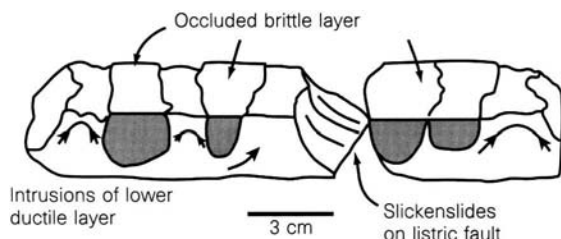
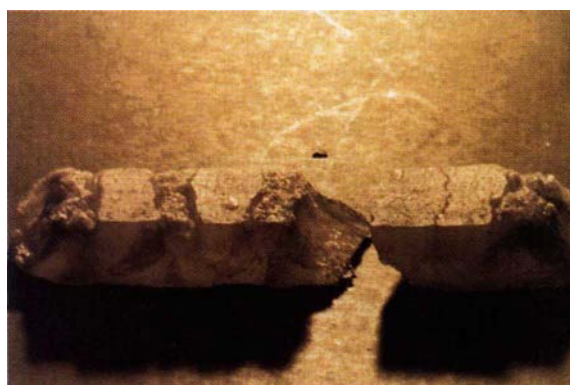


Figure 3 Photograph and line drawing of listric normal faults generated in large strain (~50%) experiments. Well-developed slickensides are evidence that shear displacement occurred across the faults. Faults begin to curve at the brittle-ductile transition, similar to the theoretical results of Melosh⁶, and root under both the brittle and, at depth, the ductile layer. The brittle layer is intruded by the lower ductile layer; see also Fig. 4 and associated text. Flow lines in ductile layer are shown by curved arrows.

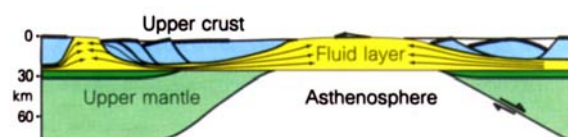


Figure 4 Top, development of oclusions of upper brittle layer as the ductile and low-viscosity layer was intruded to accommodate finite strains. (See also Fig. 3) Scale in cm is shown. We find this cross-section to be remarkably similar to a synthesis section made by Wernicke²⁰ of an extending continental lithosphere (bottom, modified from ref. 20). Strain distribution in both the wax model and cross-section is strongly inhomogeneous with depth, a critical feature of current models of lithospheric extension^{7,20,32}.

normal faults, using either elastic or viscoelastic constitutive relations, have predicted a listric fault geometry based on the orientation of principal stresses and the Mohr–Coulomb failure criterion^{4–6,12}. The appropriate stress conditions arise, in elastic models^{4,5}, from the non-zero shear stress boundary condition along the base of the crust and, in viscoelastic models, from the refraction of stress across significant rheological contrasts¹² or from rotated stresses induced by flow in the low-viscosity layer⁶. Our results are consistent with the latter, where the listric fault geometry of the fault appears to be controlled by high local internal shear strains associated with lateral flow of the lower, viscous layer.

A number of other explanations have been offered for the generation of low-angle listric normal faults. These include initiation of normal faults at high angles followed by finite rotation to low angles^{13–15}, and the presence of elevated pore-fluid pressures in a confined fault zone¹⁶ or within units of differing rheologies¹². These plausible explanations are probably limited in practice. The former is inconsistent with many observations of an initial low-angle fault dip (see, for example, refs 11, 17, 18) and the latter requires relatively special conditions to be maintained within the fault zone or bedded units. In contrast, a brittle–ductile transition is generally considered to exist everywhere at one length scale or another.

Finite deformation (~50%) of a relatively thick brittle layer (~25%) in this experiment also produced relatively large-scale intrusion of the brittle layer by material of the ductile layer, isolating blocks of the upper brittle layer (Fig. 4). The occluded brittle material was in places detached from the lower ductile material by a listric normal fault or thin shear zone (Fig. 4 top). These structures are remarkably similar to core complexes (ref. 19 and references therein) and inferred structures in the Basin and Range province (Fig. 4 bottom) and to the notion that some part of the lower or middle crust behaves as a fluid²⁰. A weak lower crust is supported by a wealth of geophysical^{21,22}, geological^{20,23} and theoretical studies^{24–26}

in (and of) mountain belts.

The development of core complexes in the wax model was most easily recognized when the brittle layer was relatively thick, but the key appears to have been the presence also of a significant thickness of the viscous layer. This is consistent with the observation that core complexes in the Basin and Range province developed when the young, hot Farallon plate was subducting underneath North America²⁷, thus providing a constant heat source to the lower crust. Emplacement of core complexes in the wax model may resemble the triggering of salt diapirism in regions of extension to the extent that extension, rather than density contrast, provides the driving force by generating a zone of low pressure into which diapirs of viscous material move²⁸.

Buck⁷ developed a model of continental lithospheric extension in which modes of extension (core complex formation, wide rift extension, and narrow rift extension) reflected initial conditions. This is similar to our findings, in that the mode of extension appears to reflect the proportion of brittle to ductile layering, a function in the real Earth of heat flow, crustal age and crustal thickness. The wax model shows, moreover, that some of the more controversial structures (listric faults, core complexes, and faults rooted under occluded upper crust) may be realized in a physical model and are thus physically likely in the Earth's crust. □

Received 25 October 1996; accepted 4 March 1997.

1. Coward, M. P., Dewey, J. F. & Hancock, P. L. *Continental Extensional Tectonics* (Spec. Publ. 28, Geol. Soc. London, 1987).
2. Oldenburg, D. W. & Brune, J. N. Ridge transform fault spreading pattern in freezing wax. *Science* **178**, 301–304 (1972).
3. Oldenburg, D. W. & Brune, J. N. An explanation for the orthogonality of ocean ridges and transform faults. *J. Geophys. Res.* **80**, 2575–2585 (1975).
4. Yin, A. Origin of regional, rooted low-angle normal faults: a mechanical model and its tectonic implications. *Tectonics* **8**, 469–482 (1989).
5. Spencer, J. E. & Chase, C. G. Role of crustal flexure in initiation of low-angle normal faults and implications for structural evolution of the Basin and Range province. *J. Geophys. Res.* **94**, 1765–1775 (1989).

6. Melosh, H. J. Mechanical basis for low-angle normal faulting in the Basin and Range province. *Nature* **343**, 331–335 (1990).
7. Buck, W. R. Modes of continental lithospheric extension. *J. Geophys. Res.* **96**, 20161–20178 (1991).
8. Longwell, C. R. Low-angle normal faults in the Basin and Range province. *Eos* **26**, 107–118 (1945).
9. Wernicke, B. & Burchfiel, B. C. Modes of extensional tectonics. *J. Struct. Geol.* **4**, 105–115 (1982).
10. Anders, M. H. & Christie-Blick, N. Is the Sevier Desert reflection of west-central Utah a normal fault? *Geology* **22**, 771–774 (1994).
11. Wernicke, B. Low-angle normal faults and seismicity: a review. *J. Geophys. Res.* **100**, 20159–20174 (1995).
12. Bradshaw, G. A. & Zoback, M. D. Listric normal faulting, stress refraction, and the state of stress in the Gulf Coast basin. *Geology* **16**, 271–274 (1988).
13. Proffett, J. M. J. Cenozoic geology of the Yerington district, Nevada, and implications for the nature and origin of Basin and Range faulting. *Geol. Soc. Am. Bull.* **88**, 247–266 (1977).
14. Vendeville, B. & Cobbold, P. R. How normal faulting and sedimentation interact to produce listric fault profiles and stratigraphic wedges. *J. Struct. Geol.* **10**, 649–659 (1988).
15. Brun, J.-P., Sokoutis, D. & Van Den Driessche, J. Analog modeling of detachment fault systems and core complexes. *Geology* **22**, 319–322 (1994).
16. Axen, G. J. Pore pressure, stress increase, and fault weakening in low-angle normal faulting. *J. Geophys. Res.* **97**, 8979–8991 (1992).
17. Davis, G. A. & Lister, G. S. in *Processes in Continental Lithospheric Deformation* (ed. Clark, S. P. J.) 133–159 (Spec. Pap. 218, Geol. Soc. Am., Boulder, 1988).
18. Scott, R. J. & Lister, G. S. Detachment faults: evidence for a low-angle origin. *Geology* **20**, 833–836 (1992).
19. Crittenden, M. D., Coney, P. J. Jr & Davis, G. H. *Cordilleran Metamorphic Core Complexes* (Mem. 153, Geol. Soc. Am., Boulder, 1980).
20. Wernicke, B. in *Exposed Cross-sections of the Continental Crust* (eds Salisbury, M. H. & Fountain, D. M.) 509–544 (Kluwer Academic, Boston, MA, 1990).
21. Hauser, E. et al. Crustal structure of eastern Nevada from COCORP deep seismic reflection data. *Geol. Soc. Am. Bull.* **99**, 833–844 (1987).
22. Kruse, S., McNutt, M., Phipps-Morgan, J. & Royden, L. Lithospheric extension near Lake Mead, Nevada: a model for ductile flow in the lower crust. *J. Geophys. Res.* **96**, 4435–4456 (1991).
23. Gans, P. B. An open-system, two-layer crustal stretching model for the eastern Great Basin. *Tectonics* **6**, 1–12 (1987).
24. Block, L. & Royden, L. H. Core complex geometries and regional scale flow in the lower crust. *Tectonics* **9**, 557–567 (1990).
25. Bird, P. Lateral extrusion of lower crust from under high topography, in the isostatic limit. *J. Geophys. Res.* **96**, 10275–10286 (1991).
26. King, G. & Ellis, M. The origin of large local uplift in extensional regions. *Nature* **348**, 689–693 (1990).
27. Coney, P. J. in *Cenozoic Tectonics and Regional Geophysics of the Western Cordillera* (eds Smith, R. B. & Eaton, G. P.) 33–50 (Geol. Soc. Am., Boulder, CO, 1978).
28. Jackson, M. P. A. & Vendeville, B. Regional extension as a geologic trigger for diapirism. *Geol. Soc. Am. Bull.* **106**, 57–73 (1994).
29. Mancktelow, N. S. The rheology of paraffin wax and its usefulness as an analogue for rocks. *Geol. Inst. Univ. Uppsala* **14**, 181–193 (1988).
30. McKenzie, D. & Bickle, M. J. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.* **29**, 625–679 (1988).
31. Hildenbrand, T. G. Rift structure of the northern Mississippi embayment from the analysis of gravity and magnetic data. *J. Geophys. Res.* **90**, 12607–12622 (1985).
32. Jones, C. H., Kanamori, H. & Roecker, S. W. Missing roots and mantle 'drips': regional Pn and teleseismic arrival times in the southern Sierra Nevada and vicinity. *J. Geophys. Res.* **99**, 4567–4601 (1994).

Acknowledgements. We thank R. Buck for comments and R. Schweickert for advice and logistical support.

Correspondence and requests for materials should be addressed to M.A.E. (e-mail: ellis@ceri.memphis.edu).

Adaptive differentiation following experimental island colonization in *Anolis* lizards

Jonathan B. Losos*, Kenneth I. Warheit† & Thomas W. Schoener‡

* Department of Biology, Campus Box 1137, Washington University, St Louis, Missouri 63130-4899, USA

† Department of Fish and Wildlife, 600 Capitol Way North, Olympia, Washington 98501-1091 and Burke Museum, DB-10, University of Washington, Seattle, Washington 98195, USA

‡ Section of Evolution and Ecology and Division of Environmental Studies, University of California, Davis, California 95616, USA

If colonizing populations are displaced into an environment that is often very different from that of their source, they are particularly likely to diverge evolutionarily, the more so because they are usually small and thus likely to change by genetic restructuring or drift^{2,3}. Despite its fundamental importance, the consequence of colonization for traits of founding popula-

tions have primarily been surmised from static present-day distributions^{1,2,4,5}, laboratory experiments⁶ and the outcomes of haphazard human introductions^{7–9}, rather than from replicated field experiments. Here we report long-term results of just such an experimental study. Populations of the lizard *Anolis sagrei*, introduced onto small islands from a nearby source, differentiated from each other rapidly over a 10–14-year period. The more different the recipient island's vegetation from that of the source, the greater the magnitude of differentiation. Further, the direction of differentiation followed an expectation based on the evolutionary diversification of insular *Anolis* over its entire geographic range. In addition to providing a glimpse of adaptive dynamics in one of the most extensive generic radiations on earth, the results lend support to the general argument that environment determines the evolution of morphology.

Propagules of 5 or 10 lizards (2:3 male:female ratio) were introduced onto 14 small islands near Staniel Cay, Exumas, Bahamas, in 1977 (11 islands)¹⁰ and 1981 (3 islands). Lizards were collected from Staniel Cay and randomly assigned to islands that did not contain lizards naturally (probably because of periodic hurricanes). On all but some of the smaller islands, lizard populations have persisted and, on some islands, thrived; one island attained a population of over 700 lizards. Consequently, these introductions constitute a replicated experiment of the effect of island characteristics and founder propagule size on differentiation among colonizing populations. The source of the introductions, Staniel Cay (roughly 3 × 1 km), is covered by a variety of vegetation including substantial amounts of coppice, a scrubby to moderately tall forest formation¹¹. Staniel contrasts markedly with the experimental islands, which have few trees and are mostly covered by narrow-diameter vegetation.

The extensive research on the evolutionary radiation of *Anolis* lizards in the Caribbean provides the framework within which these experiments were conducted. On the basis of macroevolutionary and among-population studies of *Anolis* adaptation to structural habitat, we can formulate *a priori* predictions about how introduced populations should adapt to the new structural habitats on the experimental islands. Comparisons among and within species of *Anolis* indicate that hindlimb length evolves in response to changes in the mean diameter of perches, producing a positive correlation between limb length and utilized perch diameter¹²; functional studies indicate that this correlation is probably the result of a trade-off between maximal locomotor abilities increasing with limb length versus the ability to move efficiently on narrow surfaces decreasing with limb length¹³. These macroevolutionary patterns are sufficient to make the following predictions about how anoles adapt to changes in vegetational structure: (1) the degree to which experimental populations diverge morphologically from the source population should correlate with the degree to which the vegetation structure on an experimental island differs from that on Staniel Cay; and (2) among populations, a positive relationship should exist between relative hindlimb length and mean perch diameter.

In May 1991, adult male lizards were captured and morphological measurements taken on most of the experimental islands still having lizards. In addition, 41 male *A. sagrei* were captured from the areas on Staniel Cay that the lizards for the original introductions came from. We removed the effects of size from morphological variables (see Methods) to examine differences in shape of lizards from each of the experimental islands and Staniel Cay. Populations differed in shape (Table 1, columns 3 and 4, bottom), but not body size (Table 1, column 1). Figure 1a plots the position of the populations on Staniel Cay and the experimental islands in a morphological space defined by the first two shape axes (that is, from principal components analysis on size-adjusted variables). Although the Staniel individuals were randomly distributed about their population centroid (the mean value for the two axes), with no uni- or bimodal pattern (Rayleigh's test of circular distribution,

$z = 0.70$, $P > 0.50$), the experimental islands have clearly differentiated non-randomly with respect to the source population ($z = 7.12$, $P < 0.001$); most of the experimental island centroids lie in a 90° sector relative to the Staniel centroid. The position of experimental island centroids in this two-dimensional morphological space indicates that lizards on the experimental islands had relatively shorter limbs, and, to a lesser extent, wider toe-pads and greater mass than the lizards from Staniel Cay (Fig. 1b).

Both of our hypotheses about how morphology should diverge in relation to habitat structure were confirmed. First, a relationship existed between how different the vegetation on an experimental island was from the source island and how much the population on that experimental island diverged morphologically from the source

population ($r = 0.55$, $P < 0.025$, one-tailed, Fig. 2). Analysis of covariance with vegetational difference as the covariate gave no significant effect of propagule size ($P > 0.25$). Second, a positive relationship also existed between size-adjusted hindlimb length and perch diameter used (including Staniel: $r = 0.59$, $P < 0.025$, Fig. 3; not including Staniel: $r = 0.58$, $P < 0.025$; both comparisons one-tailed).

Although rates of evolution as rapid as observed in this study are not uncommon in introduced populations¹⁴, rarely has the adaptive nature of this change been so clear-cut. In this case, the extensive knowledge of *Anolis* natural history and evolution allows *a priori* prediction of patterns of differentiation among populations introduced to an environment markedly different from the environment

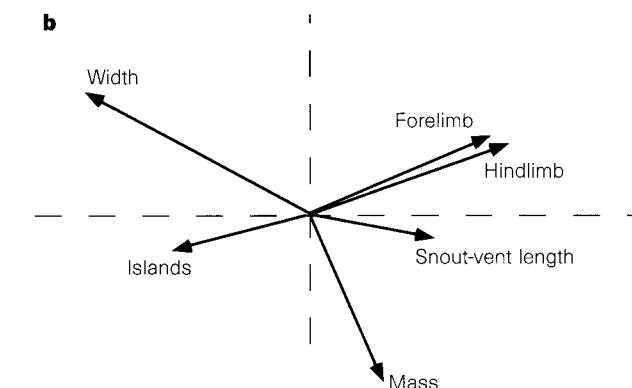
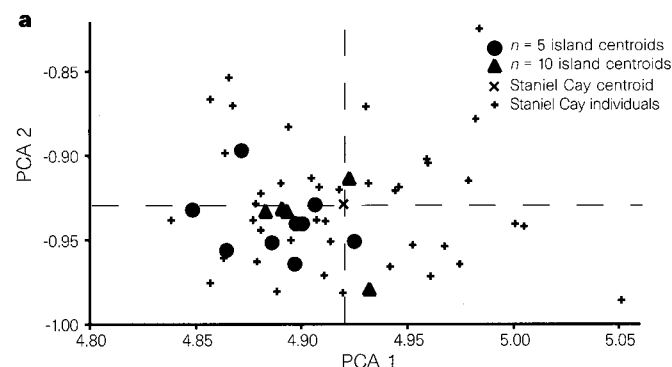


Figure 1 a, Morphological space defined by the first two size-adjusted principal components axes (PCA). Circles are the centroids for islands founded with propagules of five individuals, triangles are the centroids for islands founded with propagules of 10 individuals, the x is the centroid for the source population (Staniel), and plusses (+) are individuals from Staniel. Rate of evolution was calculated separately for the two principal component axes and ranged from 0–955 darwins for axis 1 and 0–2,117 darwins for axis 2. **b**, All vectors, except 'islands', indicate the direction and loading of each variable, ln-transformed, on the two principal components axes from the size-adjusted analyses. For example, hindlimb length loads strongly and positively on PC1 and weakly and positively on PC 2. These vectors indicate, for example, that most populations have moved in the opposite direction from the limb vectors, indicating shorter limbs. The vector labelled 'islands' marks the direction of the mean angle for the centroids of the experimental islands. This is a directional vector only (that is, length of the vector is arbitrary).

Table 1 Differentiation in size and shape among experimental populations

Variables	Pooled within-island analysis		Size-adjusted analysis	
	Axis 1*	Shape coefficient†	Axis 1	Axis 2
Snout-vent length	0.92	0.83	0.55	–0.11
Forelimb length	0.80	0.64	0.81	0.43
Hindlimb length	0.77	0.60	0.86	0.38
Lamellae width	0.94	1.56	–0.96	0.29
Mass	0.90	1.05	0.33	–0.92
Variance (%)	80.9		59.9	26.5
With Staniel‡				
	$F = 0.549$		2.757	2.754
	$P = 0.88$		0.002	0.002
Without Staniel‡				
	$F = 0.632$		2.398	3.967
	$P = 0.80$		0.01	0.0001

Size accounted for 81% of the pooled within-island variance. The loadings were large and positive for all dimensions, which indicates that the correlation between each dimension and size was large (first column).

*Size (see text).

†Multivariate allometric coefficients for the slope of the regression line relating each variable to the first principal component.

‡Results of analysis of variance.

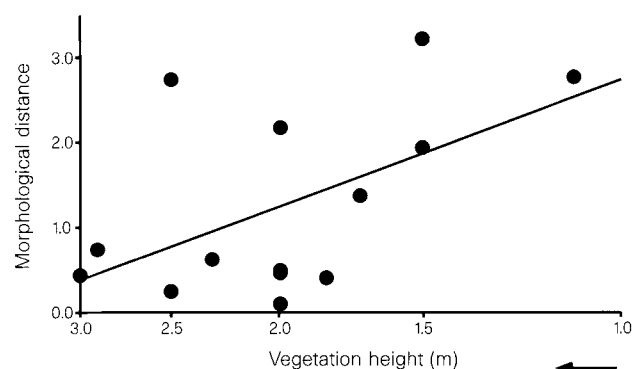


Figure 2 Vegetation height of experimental islands versus the morphological (Mahalanobis) distance of the population on that island from the Staniel population. The scale of the x-axis (from large to small heights) is reversed to emphasize that islands with low vegetation are more dissimilar from Staniel than are islands with high vegetation. Maximum vegetation height on each island was used as a measure of vegetational differences among islands; on similarly sized islands near Georgetown, Bahamas, which is less than 100 km from our study site, maximum height is related to a variety of measures of vegetation structure, such as mean vegetation height ($r = 0.82$, $P < 0.0005$, $n = 14$) and mean vegetation diameter at a height of 12 ($r = 0.53$, $P < 0.05$, $n = 14$) and 36 inches ($r = 0.62$, $P < 0.025$, $n = 12$ (two islands had too little vegetation tall enough to measure at 36 inches); all comparisons one-tailed; J.B.L. and D. Spiller, unpublished results). Staniel Cay, the source island, has vegetation considerably broader and higher than on any of the experimental islands, with some trees greater than 10 m tall. Differences in vegetation height among islands may also correlate with other variables affecting these lizards, such as available thermal microclimates.

of the source island. Further, the substantial replication of these introductions allows an examination of the concordance between the degree to which the new environment is different from the source environment and the extent to which experimental populations have differentiated. Taken together, these results indicate that founding populations of *A. sagrei*, despite their small initial size, can survive and rapidly adapt to the new environmental conditions in their new habitat.

An important parallel exists between patterns of adaptive differentiation reported here for one species of *Anolis* and patterns apparent among species of *Anolis* in the entire Caribbean *Anolis* radiation, which has produced nearly 150 species (250 more occur in Central and South America). In both cases, differences in limb dimensions are related to differences in microhabitat structure, although the differences are considerably greater among Greater Antillean species^{12,15}. This study thus indicates that not only can populations rapidly respond to new environmental conditions, but also that the response is in some ways qualitatively similar to large-scale patterns manifest on macroevolutionary timescales. This suggests that the processes operating during adaptive radiation may be similar to those producing microevolutionary adaptation—macroevolution may just be microevolution writ large—and, consequently, that insight into the former may result from study of the latter.

However, another mechanism could also produce divergence among our experimental populations similar to that seen in macroevolutionary time. In recent years, awareness of the adaptive importance of non-genetic environmental effects on morphological size and shape of animals has grown (for example refs 16, 17). A variety of environmental differences among experimental islands could potentially lead to morphological differences. The most plausible of these is differences in the diameter of vegetation used by the lizards. The different bone stresses produced by living on surfaces of different diameters could plausibly cause the limbs of lizards to grow at different rates. However, most studies (for example, refs 18, 19) of the effect of physical exertion on limb growth (all on endotherms) generally have noted differences in bone density or diameter, rather than in length, although there have been some exceptions (such as ref. 20). If this developmental hypothesis is correct, then the evolution of the large morphological differences that characterize the *Anolis* adaptive radiation may be the outcome of selection operating on a trait with an initially large environmental component. This would indicate that phenotypic plasticity may have macroevolutionary significance, as proposed in a hypothesis formulated nearly 50 years ago^{21–24}. Further observation and experimentation will be necessary to determine whether this is the case or whether the purely evolutionary hypothesis is correct. □

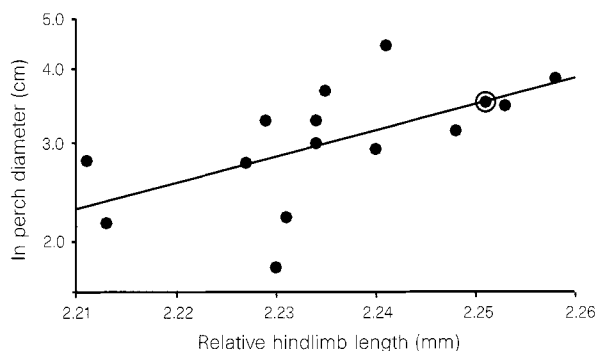


Figure 3 Relationship between mean perch diameter (ln-transformed) used and size-adjusted hindlimb length in populations of the lizard *Anolis sagrei* on islands near Staniel Cay, Bahamas. Circled point indicates Staniel. The rate of evolution of size adjusted hindlimb length varied from 89–1,195 darwins.

Methods

Data collection. The following measurements were taken on each lizard (sample size on the experimental islands = 1–14; $\bar{x}$ = 8.6): snout-vent length (svl), mass, fore- and hindlimb length, and width of the subdigital pad on the fourth hindtoe (lamella width). No individuals from the source population were measured or preserved at the time of the introductions because the experiment was designed to test very different hypotheses from those examined here¹⁰. Habitat use by the lizards was measured by recording the height and perch diameter for each male lizard observed (sample size = 7–57; $\bar{x}$ = 26.7). The range in vegetated area occupied by lizards for experimental islands was 89–5,790 m². Although island species turnover is much lower for lizards than many other kinds of organisms²⁵, it does occur. Thus, dispersal between naturally occupied and experimental populations is possible; its precise importance remains to be determined by future theoretical and empirical evaluation.

Multivariate analysis of size differentiation. We define size as the first principal component from a pooled within-island matrix of morphological variables, thus following rubric 2 in ref. 26. The pooled within-group matrix is equal to the error mean square matrix from a MANOVA^{27,28}. We justified the use of this multivariate size dimension, rather than a univariate size dimension, because each of the five morphological variables increases at different rates with lizard growth, and only the adjusted mass dimension approached isometry (Table 1, column 2). Because variance–covariance matrices require that all variables be measured in the same units, we converted mass to a linear dimension using the equation²⁹ adjusted mass (mm) = $35 \times \text{mass}^{0.34}$. Thus, no one dimension best represents size because each dimension represents growth in a different way. Therefore, a multivariate size dimension that (1) uses information from all univariate dimensions, (2) accounts for the variance–covariance structure of the data set and (3) quantifies allometry, best characterizes growth or ontogenetic size (other methods (such as ref. 12) produce qualitatively similar results).

Multivariate shape analyses. We removed the effects of size from the original log-transformed data set using the method of Burnaby³⁰. A principal components analysis using a total variance–covariance matrix was then conducted on these size-adjusted variables. This produced composite shape variables based on the covariance structure of a size-free data set. To quantify the extent to which the populations diverged from the source population, we calculated the Mahalanobis distance in morphometric space, defined by the first two size-adjusted principal component axes, from each experimental population to the Staniel population. Populations on islands 13 and 25¹⁰ could not be included in the Burnaby analysis because of their small sample sizes; for these populations, size-adjusted variables were calculated by projecting the ln-transformed variables from each captured lizard from these two islands onto the original Burnaby size-adjusted plane defined using the other islands. Mahalanobis distances were estimated for the populations on these two islands using the pooled variance–covariance structure of the other islands.

Rates of evolution. Rates of morphological evolution were calculated as $r = (\ln(x_2) - \ln(x_1))/\Delta t$, where r = rate of change (in darwins), x_1 and x_2 are the initial and final dimensions of the character, and Δt is the time elapsed in millions of years. These rates were compared to those calculated for a variety of other recently introduced species¹⁴.

Received 27 November 1996; accepted 3 March 1997.

1. Carlquist, S. *Island Biology* (Columbia Univ. Press, New York, 1974).
2. Carson, H. L. & Templeton, A. R. Genetic revolutions in relation to speciation phenomena: the founding of new populations. *Ann. Rev. Ecol. Syst.* **15**, 97–131 (1984).
3. Templeton, A. R. The theory of speciation via the founder principle. *Genetics* **94**, 1011–1038 (1980).
4. Darlington, P. J. *Zoogeography: the Geographical Distribution of Animals* (Wiley, New York, 1957).
5. Williams, E. E. The ecology of colonization as seen in the zoogeography of anoline lizards on small islands. *Q. Rev. Biol.* **44**, 345–389 (1969).
6. Templeton, A. R. Experimental evidence for the genetic-transience model of speciation. *Evolution* **50**, 909–915 (1996).
7. Baker, A. J. Genetic and morphometric divergence in ancestral European and descendant New Zealand populations of chaffinches (*Fringilla coelebs*). *Evolution* **46**, 1794–1800 (1992).
8. St. Louis, V. L. & Barlow, J. C. Morphometric analyses of introduced and ancestral populations of the Eurasian tree sparrow. *Wilson Bull.* **103**, 1–12 (1991).
9. Johnston, R. F. & Selander, R. K. Evolution in the house sparrow. II. Adaptive differentiation in North American populations. *Evolution* **25**, 1–28 (1971).
10. Schoener, T. W. & Schoener, A. The time to extinction of a colonizing propagule of lizards increases with island area. *Nature* **302**, 332–334 (1983).
11. Coker, W. C. in *The Bahama Islands* (ed. Shattuck, G. B.) 185–270 (MacMillan, New York, 1905).
12. Losos, J. B., Irschick, D. J. & Schoener, T. W. Adaptation and constraint in the evolution of specialization of Bahamian *Anolis* lizards. *Evolution* **48**, 1786–1798 (1994).

13. Losos, J. B. & Irschick, D. J. The effect of perch diameter on escape behaviour of *Anolis* lizards: laboratory predictions and field tests. *Anim. Behav.* **51**, 593–602 (1996).
14. Gingerich, P. D. Rates of evolution: effects of time and temporal scaling. *Science* **222**, 159–161 (1983).
15. Williams, E. E. in *Lizard Ecology: Studies of a Model Organism* (eds Huey, R. B., Pianka, E. R. & Schoener, T. W.) 326–370 (Harvard Univ. Press, Cambridge, MA, 1983).
16. Patton, J. L. & Brylski, P. V. Pocket gophers in alfalfa fields: causes and consequences of habitat-related body size variation. *Am. Nat.* **130**, 493–506 (1988).
17. Robinson, B. W. & Wilson, D. S. Experimentally induced morphological diversity in Trinidadian guppies (*Poecilia reticulata*). *Copeia* **1995**, 294–305 (1995).
18. Woo, S. L.-Y. et al. The effect of prolonged physical training on the properties of long bone: a study of Wolff's Law. *J. Bone Joint Surg.* **63-A**, 780–786 (1981).
19. Loitz, B. J. & Zernicke, R. F. Strenuous exercise-induced remodelling of mature bone: relationships between in vivo strains and bone mechanics. *J. Exp. Biol.* **170**, 1–18 (1992).
20. Kiiskinen, A. Physical training and connective tissues in young mice: physical properties of achilles tendons and long bones. *Growth* **41**, 123–137 (1977).
21. Schmalhausen, I. I. *Factors of Evolution* (Blakiston, Philadelphia, PA, 1949).
22. Huxley, J. *Evolution: the Modern Synthesis* (Harper, New York, 1942).
23. Waddington, C. H. *The Evolution of an Evolutionist* (Columbia Univ. Press, New York, 1975).
24. West-Eberhard, M. J. Phenotypic plasticity and the origins of diversity. *Ann. Rev. Ecol. Syst.* **20**, 249–278 (1989).
25. Schoener, T. W. in *Community Ecology* (eds Diamond, J. & Case, T. J.) 556–586 (Harper and Row, New York, 1986).
26. Bookstein, F. L. "Size and shape": a comment on semantics. *Syst. Zool.* **38**, 173–180 (1989).
27. Bookstein, F. L. et al. (eds) *Morphometrics in Evolutionary Biology* (Acad. Natl. Sci., Philadelphia, PA, 1985).
28. Rohlf, F. J. & Bookstein, F. L. A comment on shearing as a method of "size correction". *Syst. Zool.* **36**, 356–367 (1987).
29. Stamps, J. A., Krishnan, V. V. & Andrews, R. M. Analyses of sexual size dimorphism using null growth-based models. *Copeia* **1994**, 598–613 (1994).
30. Burnaby, T. P. Growth-invariant discriminant functions and generalized distances. *Biometrics* **22**, 96–110 (1966).

Acknowledgements. We thank M. Krukowski, R. Lewis, S. Rissing, S. Rohwer, J. Smith, D. Spiller, S. Sultan, S. Swartz and the 1993 Burke Museum discussion group for helpful assistance and suggestions and the NSF for support. We also thank the Bahamian government for permission to do this work.

Correspondence and requests for materials to J.B.L. (Losos@biodec.wustl.edu).

Contrast dependence of contextual effects in primate visual cortex

Jonathan B. Levitt & Jennifer S. Lund

Department of Visual Science, Institute of Ophthalmology, University College London, Bath Street, London EC1V 9EL, UK

The responses of neurons in the visual cortex to stimuli presented within their receptive fields can be markedly modulated by stimuli presented in surrounding regions that do not themselves evoke responses^{1–7}. This modulation depends on the relative orientation and direction of motion of the centre and surround stimuli, and it has been suggested that local cortical circuits linking cells with similar stimulus selectivities underlie these phenomena^{8–16}. However, the functional relevance and nature of these integrative processes remain unclear. Here we investigate how such integration depends on the relative activity levels of neurons at different points across the cortex by varying the relative contrast of stimuli over the receptive field and surrounding regions. We show that simply altering the balance of the excitation driving centre and surround regions can dramatically change the sign and stimulus selectivity of these contextual effects. Thus, the way that single neurons integrate information across the visual field depends not only on the precise form of stimuli at different locations, but also crucially on their relative contrasts. We suggest that these effects reflect a complex gain-control mechanism that regulates cortical neuron responsiveness, which permits dynamic modification of response properties of cortical neurons.

Figure 1 shows the responses of five representative V1 neurons, A–E. Using high-contrast stimuli confined to the minimum response field (centre), we first determined each neuron's orientation and direction selectivity. Here the centre stimulus was surrounded by a uniform field with the same mean luminance. We then paired the optimal central stimulus with surround stimuli of

varying orientation and drift direction. The third column in Fig. 1 shows that pairing the optimal high-contrast centre stimulus with a surround stimulus matching the centre in both orientation and drift direction often elicited a profound response suppression (relative to the centre-alone condition). When the surround stimulus differed in orientation or direction of motion from that in the centre, the response suppression often disappeared. Responses then equalled or even exceeded those obtained under control conditions using the central stimulus alone. Although response suppression was most common when centre and surround stimuli had matching orientations, we did observe some cases where oblique or orthogonal surround orientations were most suppressive. In these cases, suppression again disappeared as the orientation of the surround stimulus varied from that configuration. This behaviour matches previously described stimulus-specific response modulation from beyond the classical receptive field^{1–7}.

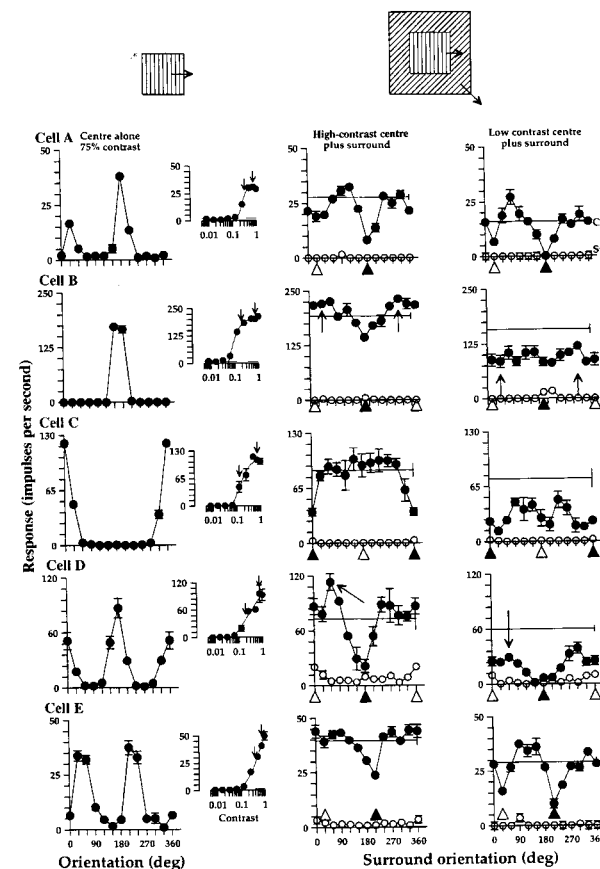


Figure 1 Representative V1 neuron responses. Stimulus configurations are shown schematically at the top. Each row (A–E) illustrates the mean responses ($\pm$ s.e.m.) of a different cell. Left column: responses to grating stimuli of varying orientations (contrast 0.75) confined to the cell's minimum response field. Second column: contrast response functions (using optimal centre stimulus alone) to determine subsequent choice of centre contrasts (arrows). Contrasts of surround and high contrast centre stimuli were always 0.75. Low-centre contrasts were chosen to be near the middle of the cell's dynamic range (A: layer 4A, 0.32; B: layer 6, 0.185; C: layer 4B, 0.125; D: layer 3, 0.185; E: layer 3, 0.38). Two right-hand columns: responses to the optimal central stimulus (high and low centre contrast) paired with surround gratings of varying orientation (filled circles), and to the surround stimuli alone (circles). Horizontal line indicates response to the optimal centre stimulus alone. Filled triangles indicate the fixed centre stimulus orientation, open triangles indicate the same orientation stimulus drifting in the opposite direction. Arrows indicate identical stimulus conditions (differing only in the relative contrast of centre and surround) with different effects at high and low centre contrast.

In many cases (cell A), the stimulus selectivity of this modulation did not depend on centre stimulus contrast. However, in many other cases we did observe striking changes when surround stimuli were paired with lower-contrast centre stimuli (shown in the right column of Fig. 1). Although we still often observed profound response suppression when the surround stimulus matched that in the centre, the modulatory effects exerted by other surround stimuli were quite different. The first notable feature was a change in the tuning of modulatory effects. At high centre stimulus contrast, response suppression was often limited to the configuration where central and surround stimuli were identical. Comparison of the two right-hand columns in Fig. 1 shows that for some cells, more stimuli suppressed neuronal responses at low centre contrast than at high centre contrast. However, this resulted in a change in the tuning of suppression, not merely in its amplitude. Using a high centre contrast, cell B of Fig. 1 showed response suppression by surround stimuli matching in orientation and drift direction, and facilitation by surrounds moving in orthogonal or opposite directions. At low centre contrast, however, its responses were suppressed by all surround stimuli. Cells C, D and E are further examples of cells suppressed most at high centre contrast by surround stimuli matching the centre in orientation and drift direction, whereas at

low centre contrast they were also suppressed by surrounds of matching orientation drifting in the opposite direction (indicated by open triangles). In some cases (cell E), this resulted from the selective additional suppression of stimuli matching in orientation but moving in the opposite direction, whereas in other cells (cells C, D), this was accompanied by suppression at all other surround orientations. Thus for some V1 neurons, the suppression originating from the receptive field surround changed from direction selective to non-direction selective (or completely non-selective) merely by changing the contrast of the central stimulus.

Figure 2 replots these data to express responses in both centre contrast conditions as the difference from the response in the centre alone condition. For cells like B–E, the orientation and direction dependence of surround modulation differed dramatically in the two centre contrast conditions, and cannot be explained simply as a vertical shift in the suppression tuning curves because the shapes of the curves differ. In other words, these changes in the tuning of suppression did not result simply as a consequence of a change in the amplitude of suppression. We quantified the change in tuning of suppression by computing an orientation suppression index (OSI) and a directional suppression index (DSI) for each cell in both high and low centre contrast conditions (see Methods). The orientation

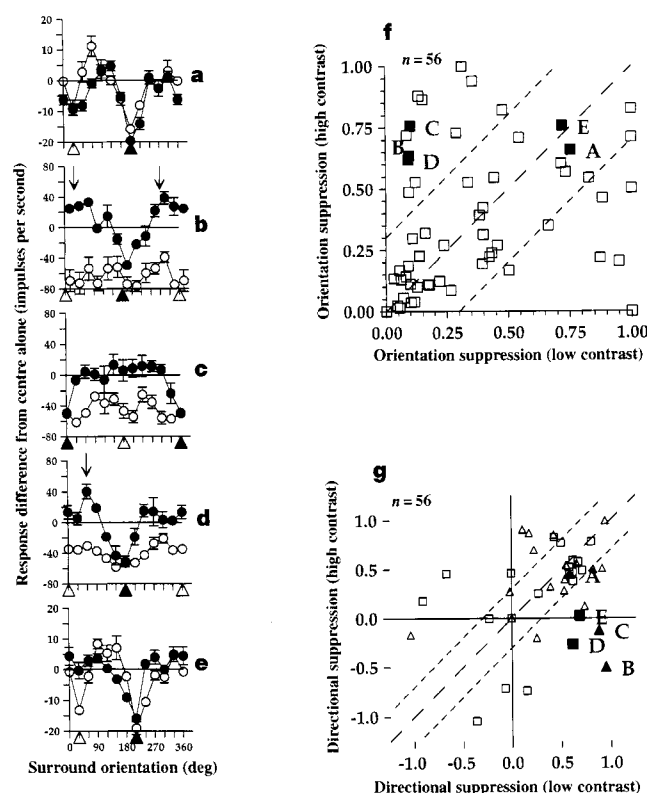


Figure 2 a–e. Responses of Fig. 1 replotted to indicate how responses to the centre-surround combinations differed from the centre alone condition in each case. Negative numbers indicate suppression, positive numbers indicate facilitation, and zero indicates no difference from the centre alone conditions, using high (filled circles) and low (open circles) centre contrasts. Arrows again indicate identical stimulus configurations that led to either facilitation or suppression, depending on centre stimulus contrast. **f, g.** Scatterplots of the orientation suppression index and directional suppression index for each cell measured at low and high centre contrasts. Cells A–E are indicated by filled symbols. In **g**, triangles represent direction-selective neurons, and square symbols represent nondirectional neurons. Fine dashed lines indicate the range within which these indices did not vary with centre contrast.

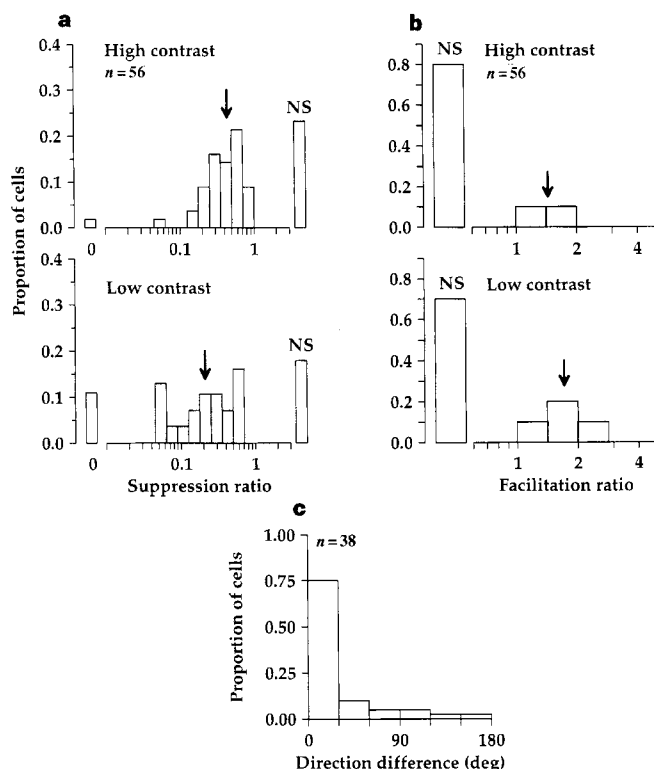


Figure 3 a, b. Histograms showing the magnitude of significant surround suppression (**a**) and facilitation (**b**) at high and low centre contrasts. For each cell, we computed a suppression ratio ($R_{\min}/R_{\text{ctr}}$) and a facilitation ratio ($R_{\max}/R_{\text{ctr}}$), where R_{ctr} is the response to centre stimulus alone, and $R_{\min}$ and $R_{\max}$ are the responses to the centre-surround combinations eliciting minimum and maximum responses. Smaller suppression ratios indicate smaller minimum responses (relative to the centre-alone condition) in the presence of the surround stimulus; larger facilitation ratios indicate greater peak responses (relative to centre alone) in the presence of the surround stimulus. Ratios close to 1 indicate no surround effect. Bins marked 'NS' indicate cells with no significant surround effects. Arrows indicate median suppression and facilitation ratios. Suppression: high contrast, 0.45; low contrast, 0.22; facilitation: high contrast, 1.44; low contrast, 1.71. **c.** Distribution of the direction differences between the maximally suppressive surround stimuli at high and low centre contrasts for those 38 cells showing significant suppression at both contrasts.

suppression index is equivalent to the circular variance of responses; indices of zero indicate that surrounds of all orientations equally suppressed cell responses, whereas indices of unity indicate that only a single surround orientation suppressed responses. Directional suppression indices near zero indicate completely direction-selective suppression (that is, suppression by surround stimuli of a particular orientation drifting in one direction of motion); directional indices near unity indicate non-directional suppression (that is, suppression by surround stimuli of a particular orientation, regardless of drift direction). Directional indices of less than zero indicate response facilitation by surround stimuli drifting in the direction of motion opposite to that causing suppression.

Figure 2f, g plots these orientation and directional suppression indices for our entire V1 sample: data points falling along the dashed diagonal line through the origin represent cells for which the orientation or direction selectivity (tuning) of surround suppression did not depend on centre contrast (regardless of changes in the amplitude of suppression). Points falling below those diagonals represent cells for which the suppression was less direction- and more orientation-selective at low centre contrasts; above the diagonals, suppression was more direction- and less orientation-selective at low centre contrast. To classify cells as differing in the two conditions, we chose a criterion OSI and DSI difference of 0.3 (namely that $R_{ctr} - R_{opp}$ must differ in the two conditions by at least 30% of $R_{ctr} - R_{min}$); this criterion is indicated in Fig. 2f, g by the finer dashed diagonal lines bracketing the dashed line through the origin. A number of V1 neurons in our sample fall outside this criterion zone. In our sample of 56 V1 neurons, 18 (32%) did not show any significant suppression in one or both centre contrast conditions; these cells appear as the cloud of data points in the lower left corner of Fig. 2f, and are all plotted at the origin in Fig. 2g as their DSI values are meaningless. Another 18 cells (32%) fell within the fine dashed diagonal lines in Fig. 2g, so the direction selectivity of surround suppression did not depend on contrast (for example, cell A). However, over a third of our sample clearly fell outside this range. In Fig. 2g, 9 cells (16%) fell above the criterion zone (suppression was more direction-selective at low centre contrast) and 11 cells (20%) fell below the criterion zone (suppression was less direction-selective at low centre contrast, as for cell C for example). Similarly, 12 cells (21%) fell above the criterion zone in Fig. 2f (suppression was less orientation selective at low centre contrast), whereas 6 cells (11%) fell below the criterion zone (suppression was more orientation selective at low centre contrast). Considering only cells showing significant surround effects, over half of all cells show this contrast dependence.

Another distinct effect was that identical stimulus configurations could be either facilitatory or suppressive, depending on the contrast of the central stimulus. In cell B (Fig. 1) for example, when the orientation of the surround stimulus differed from that of a high contrast centre by 150 deg (indicated by the arrows), the response of the cell was significantly enhanced by ~30 impulses per second. However, when the identical stimulus configuration was used with a low contrast centre stimulus, the effect of the configuration was suppressive. Similar behaviour is illustrated in cell D. This is again seen more clearly in Fig. 2. Roughly 20% (11/56) of the cells in our sample showed a shift in the sign of a surround effect with a particular surround stimulus. We emphasize that these are not simply random fluctuations in responsiveness, but rather represent significant changes in response above and below centre alone conditions. Facilitatory effects were seen in all cortical layers. We also note that facilitation of low contrast stimuli was not more prominent in cells that appeared to receive mainly magnocellular input: cell D, for example, was in layer 3 and did not have especially high contrast sensitivity.

These changes in contextual effects did not simply reflect response saturation at high centre contrast: perhaps more modulatory effects ought to be seen at low centre contrasts for that reason,

but this is not the case. Figure 3a and b shows that although there was on average a larger effect from the surround at low contrast, the proportion of cells showing no significant effect was not notably greater at high contrast, arguing against response saturation as an explanation of our results. Cells B and D in Figs 1 and 2 are clear examples of dramatic response enhancement only at high centre contrast. We note that the particular surround stimulus that produced the largest effect remained invariant with centre contrast. Figure 3c plots the direction difference in degrees between the surround stimuli producing the maximal suppressive effect for each cell at high and low centre contrasts. For most cells, the difference was less than 30 deg, indicating that the same surround stimulus exerted the greatest suppressive effect whether it was paired with a high or low contrast centre target. Thus, varying centre contrast changes the way in which other inputs are revealed, whereas the major modulatory influence seems to remain constant for a given neuron.

We observed these contrast-dependent changes in contextual effects in all cortical layers (in both simple and complex cells), suggesting that these phenomena are a fundamental feature of cortical function, not restricted to particular input or output cell populations. We have shown that the way single neurons integrate information across the visual field depends not only on the precise form of stimuli at different locations, but also on their relative contrasts (that is, on the relative activity levels of neurons at different locations across the cortex). Although the role of contextual effects remains unclear, they may mediate detection of orientation discontinuities or of weak signals in noise, or figure-ground segregation^{7,17–20}. Our result that a wider range of stimuli can modulate neuronal responses when centre and surround contrasts differ may well be consistent with this suggestion, because contrast differences could provide another cue that regions differ. However, the effects we describe here had only a weak dependence (if any) on the relative spatial phase between centre and surround. We observed strong iso-orientation suppression even in cases where centre and surround had identical contrasts and were in phase, when no distinct central region could be distinguished perceptually; this is inconsistent with a pop-out mechanism. We suggest instead that these effects reflect a complex gain control mechanism to regulate cortical responsiveness. Theoretical studies have shown the importance of recurrent excitation and inhibition to neuronal behaviour in cortex^{20–22}. Our results further suggest that under different conditions, the neuronal pool providing feedback to a given neuron can vary dynamically. The integrative biophysical properties of individual cortical neurons and the circuits among them are far more complicated than previously appreciated^{23–26}. The contrast-dependent interactions among cortical neurons we have demonstrated here highlight further complexity in the integrative processes occurring in the cerebral cortex. □

Methods

We recorded single-unit activity from parafoveal striate cortex in anaesthetized (Sufentanil: 4–8 $\mu\text{g kg}^{-1} \text{h}^{-1}$ with 50:50 O_2 and N_2O respiration) and paralysed (vecuronium bromide: 0.1 $\text{mg kg}^{-1} \text{h}^{-1}$) primate (macaque) striate cortex. All procedures conformed to NIH and Home Office guidelines. Animals viewed visual stimuli on a BARCO ICD 451B colour monitor driven by an AT Truevision Graphics board. Stimuli were drifting achromatic sinewave gratings with a mean luminance of 37.5 cd m^{-2} and subtended 13×13 deg of viewing angle. In our stimulus conventions, 0 deg represents a vertically oriented grating stimulus drifting to the right, and 180 deg the same orientation drifting to the left; stimuli of 90 and 270 deg represent horizontal gratings drifting upwards and downwards respectively. For complex cells, we used the d.c. firing rate as response measure; for simple cells we used the F1 response modulation.

We first presented small (<0.5 deg diameter) stimulus patches of optimal orientation, spatial and temporal frequency at different positions to ensure that the cell's receptive field was centred on the display screen, and to measure the cell's minimum response field. We periodically checked that the field was

centred on the TV screen during the course of studying each cell. We defined the size of each cell's receptive field centre by independently varying the width and height of stimulus patches. In such experiments, responses typically grew to a peak value, and then either became asymptotic or decreased as inhibition intervened. Centre size was defined as the stimulus diameter at which responses first reached a plateau or began to decrease. Surround stimuli were confined to a region beyond this, with an outer diameter of 13 deg; these were also sinewave grating stimuli of the same spatial and temporal frequency as the centre. Surrounds so defined usually elicited no responses themselves; occasionally we found it necessary to increase the centre diameters slightly to eliminate the residual surround response. To avoid confounding response variability with genuine response changes, we always interleaved presentations of the centre stimulus alone, the surround stimuli alone, and the two combined. Within each experimental block (individual panels in Fig. 1), we interleaved all stimuli in random order, and averaged several repeats of each stimulus condition to control for random variation in neuronal excitability. Thus, the comparisons made are of absolute response values within each block (that is, centre plus surround relative to centre alone at each centre contrast). One cannot compare absolute response values between blocks as these conditions were not interleaved, and we did not do so. Between blocks we compared only the tuning of suppression. Therefore, variations in absolute response between columns in Fig. 1 do not confound our conclusions. Relative spatial phase between centre and surround was always held constant (centre and surround aligned at field centre).

For each cell, we defined an orientation suppression index, OSI, as $|\Sigma R_i \exp(i2\Theta_i)| / \Sigma R_i$, where R_i are the response suppression values at surround orientations $0^\circ \leq \Theta_i < 180^\circ$ (response at each orientation is averaged over both drift directions). We also defined a directional suppression index, DSI as $(R_{ctr} - R_{opp}) / (R_{ctr} - R_{min})$, where R_{ctr} is the response to centre stimulus alone, R_{min} the response to the centre-surround combination eliciting the minimum response, and R_{opp} is the response to the centre-surround combination with the surround drifting in the opposite direction from that eliciting R_{min} .

Received 23 December 1996; accepted 28 February 1997.

- Nelson, J. I. & Frost, B. J. Orientation-selective inhibition from beyond the classical visual receptive field. *Brain Res.* **139**, 359–365 (1978).
- Allman, J., Miezin, F. & McGuinness, E. Stimulus specific responses from beyond the classical receptive field—neurophysiological mechanisms for local global comparisons in visual neurons. *Annu. Rev. Neurosci.* **8**, 407–430 (1985).
- Gilbert, C. D. & Wiesel, T. N. The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Visual Res.* **30**, 1689–1701 (1990).
- Knierim, J. J. & Van Essen, D. C. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *J. Neurophysiol.* **67**, 961–980 (1992).
- Li, C. Y. & Li, W. Extensive integration field beyond the classical receptive field of cat striate cortical neurons—classification and tuning properties. *Vision Res.* **34**, 2337–2355 (1994).
- DeAngelis, G. C., Freeman, R. D. & Ohzawa, I. Length and width tuning of neurons in the cat's primary visual cortex. *J. Neurophysiol.* **71**, 347–374 (1994).
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J. & Davis, J. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature* **378**, 492–496 (1995).
- Rockland, K. S. & Lund, J. L. Intrinsic laminar lattice connections in primate visual cortex. *J. Comp. Neurol.* **216**, 303–318 (1983).
- Gilbert, C. D. & Wiesel, T. N. Clustered intrinsic connections in cat visual cortex. *J. Neurosci.* **3**, 1116–1133 (1983).
- Lund, J. S., Yoshioka, T. & Levitt, J. B. Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cereb. Cortex* **3**, 148–162 (1993).
- Amir, Y., Harel, M. & Malach, R. Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *J. Comp. Neurol.* **334**, 19–46 (1993).
- Yoshioka, T., Blasdel, G. G., Levitt, J. B. & Lund, J. S. Relation between patterns of intrinsic lateral connectivity, ocular dominance, and cytochrome-oxidase reactive regions in macaque monkey striate cortex. *Cereb. Cortex* **6**, 297–310 (1996).
- Rockland, K. S., Lund, J. S. & Humphrey, A. L. Anatomical banding of intrinsic connections in striate cortex of three shrews (*Tupaia glis*). *J. Comp. Neurol.* **209**, 41–58 (1982).
- Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *J. Neurosci.* **6**, 1160–1170 (1986).
- Gilbert, C. D. & Wiesel, T. N. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J. Neurosci.* **9**, 2432–2442 (1989).
- Malach, R., Amir, Y., Harel, M. & Grinvald, A. Relationship between intrinsic connections and functional architecture revealed by optical imaging and *in vivo* targeted biocytin injections in primate striate cortex. *Proc. Natl Acad. Sci. USA* **90**, 10469–10473 (1993).
- Lamme, V. A. F. The neurophysiology of figure-ground segregation in primary visual cortex. *J. Neurosci.* **15**, 1605–1615 (1995).
- Zipser, K., Lamme, V. A. F. & Schiller, P. H. Contextual modulation in primary visual cortex. *J. Neurosci.* **16**, 7376–7389 (1996).
- Toth, L. J., Rao, S. C., Kim, D.-S., Somers, D. & Sur, M. Subthreshold facilitation and suppression in primary visual cortex revealed by intrinsic signal imaging. *Proc. Natl Acad. Sci. USA* **93**, 9869–9874 (1996).
- Douglas, R. J., Koch, C., Majowald, M., Martin, K. A. C. & Suarez, H. H. Recurrent excitation in neocortical circuits. *Science* **269**, 981–985 (1995).
- Heeger, D. J. Normalization of cell responses in cat striate cortex. *Vis. Neurosci.* **9**, 181–197 (1992).
- Somers, D. C., Nelson, S. B. & Sur, M. An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* **15**, 5448–5465 (1995).

- Stuart, G. & Sakmann, B. Amplification of EPSPs by axosomatic sodium channels in neocortical pyramidal neurons. *Neuron* **15**, 1065–1076 (1995).
- Markram, H., Helm, P. J. & Sakmann, B. Dendritic calcium transients evoked by single back-propagating action potentials in rat neocortical pyramidal neurons. *J. Physiol. (Lond.)* **485**, 1–20 (1995).
- Thomson, A. M. & Deuchars, J. D. Temporal and spatial properties of local circuits in neocortex. *Trends Neurosci.* **17**, 119–126 (1994).
- Deuchars, J. D., West, D. C. & Thomson, A. M. Relationships between morphology and physiology of pyramid–pyramid single axon connections in rat neocortex *in vitro*. *J. Physiol. (Lond.)* **478**, 423–435 (1994).

Acknowledgements. This work was supported by the NIH, the MRC and the HFSP. We thank Janssen Pharmaceuticals for the gift of sufentanil citrate, and A. Derrington, K. Gegenfurtner, D. Kiper, M. Morgan and J. Solomon for commenting on the manuscript.

Correspondence and requests for materials to J.B.L. (e-mail: j.levitt@ucl.ac.uk).

Neuronal release of soluble nucleotidases and their role in neurotransmitter inactivation

Latchezar D. Todorov*, Svetlana Mihaylova-Todorova*, Timothy D. Westfall†, Peter Sneddon†, Charles Kennedy†, Richard A. Bjur* & David P. Westfall*

* Department of Pharmacology, University of Nevada School of Medicine, Reno, Nevada 89557, USA

† Department of Physiology and Pharmacology, University of Strathclyde, Royal College, 204 George Street, Glasgow G1 1XW, UK

Efficient control of synaptic transmission requires a rapid mechanism for terminating the actions of neurotransmitters. For amino acids and monoamines, this is achieved by their uptake into the cell by specific high-affinity transporters¹; acetylcholine is first broken down in the extracellular space and then choline is taken up by the cell². Because ATP is hydrolysed to adenosine by membrane-bound enzymes (ectonucleotidases) that are present in most tissues^{3,4}, it has been assumed that these enzymes terminate the neurotransmitter actions of ATP in the brain⁵ and in the periphery^{6,7}. We show here, however, that stimulation of sympathetic nerves innervating the guinea-pig vas deferens releases not only neuronal ATP, but also soluble nucleotidases that break down this ATP to adenosine, indicating that inactivation of ATP is increased by nerve activity. This release of specific nucleotidases together with ATP represents a new mechanism for terminating the actions of a neurotransmitter.

Stimulation of the sympathetic nerves innervating the guinea-pig vas deferens evokes a biphasic contraction, which has a predominantly purinergic, rapid initial transient component and a large noradrenergic secondary tonic phase (see ref. 8 for review). Like the purinergic component of contraction, the overflow of ATP during sympathetic nerve stimulation at 8 Hz for 60 s occurred in a transient fashion. ATP appeared in the superfusate early in a train of stimuli, peaked at 20 s and then decreased rapidly (Fig. 1a, b). The products of ATP hydrolysis, ADP and AMP, showed a similar pattern of overflow, whereas the peak overflow of adenosine was slightly later than the nucleotides. Thus, endogenously released ATP is substantially degraded, with adenosine as the principal metabolite.

To study this breakdown of ATP, we examined the metabolism of exogenous $1N^6$ -etheno-ATP (e-ATP) added in the bathing solution in order to distinguish exogenously added e-ATP from endogenously released ATP. Also, using fluorescent e-ATP allowed a more sensitive detection of purine metabolism. Formation of e-ADP, e-AMP and e-adenosine from e-ATP occurs because the etheno group in e-ATP is on the adenine moiety and so fluorescent metabolites are produced by sequential dephosphorylation of e-ATP. Figure 1c, d

shows that breakdown of exogenous e-ATP (100 nM) by the tissues was minimal over 60 seconds. The amount of e-ATP in the superfusate decreased from $92 \pm 1.4\%$ to $82 \pm 0.9\%$ of the total purines. This was accompanied by an increase of e-ADP (from $9.8 \pm 0.3\%$ to $13.5 \pm 0.7\%$) and the appearance of e-adenosine ($1.8 \pm 0.2\%$ of total purines). Similar results were obtained using exogenous ATP rather than e-ATP (not shown).

This breakdown of exogenous e-ATP can be attributed to the action of plasma membrane-bound ectonucleotidases^{4,9,10}. The small extent of metabolism is consistent with previous studies on smooth muscle, in which the half-life of exogenous ATP is ~ 30 min^{4,11-13}, but contrasts with the substantial breakdown of

neuronally released ATP shown in Fig. 1a, b, suggesting that nerve stimulation somehow enhances ATP breakdown. Therefore, we followed the metabolism of exogenous e-ATP during sympathetic nerve stimulation. (Endogenously released purines were not derivatized in these experiments and so did not interfere with measurement of exogenous e-ATP and its metabolites.)

During nerve stimulation (8 Hz; 60 s), breakdown of exogenous e-ATP was much greater than in the unstimulated preparation (compare Fig. 1e, f with c, d), being almost totally hydrolysed. During the second half of the stimulation period, there was a decrease in the formation of e-adenosine and a corresponding increase in e-ADP. Despite changes in the ratio of the individual

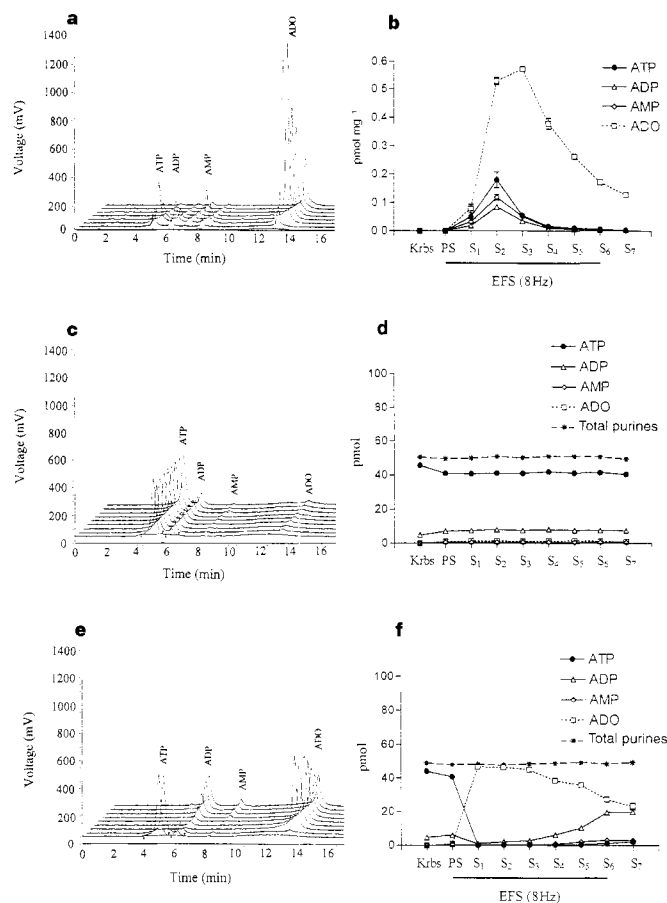


Figure 1 Release and metabolism of endogenous ATP and metabolism of exogenous e-ATP by the guinea-pig vas deferens. Left, original chromatograms for superfusate samples collected at 10-s intervals (arranged from front to back) in a single experiment. Right, converted mean data. The abscissa indicates the 10-s interval during which each sample was collected: Krebs, a control sample of Krebs solution; PS, 10 s immediately before nerve stimulation; S1-S6, 6 consecutive 10-s intervals during nerve stimulation; S7, 10 s after nerve stimulation. **a, b**, Electrical field stimulation of the sympathetic nerves (EFS; 8 Hz for 60 s, as indicated by bar) evoked a transient increase in purine overflow ($n = 8$). Adenosine was the principal purine detected, suggesting that endogenously released ATP undergoes substantial degradation within 60 s. **c, d**, Purine metabolism was studied by including 100 nM e-ATP in the bathing solution. In the absence of nerve stimulation there was little breakdown of e-ATP over 60 s ($n = 5$). Note that S1-S7 represent seven consecutive 10-s samples collected in the absence of nerve stimulation. **e, f**, In contrast, e-ATP was almost entirely degraded if the sympathetic nerves were concomitantly stimulated at 8 Hz for 60 s ($n = 5$). Adenosine was the principal metabolite, similar to the data shown in **a** and **b** for endogenously released ATP. Standard errors of the mean are shown in all figures; when it is not visible, the standard error falls within the size of the symbol.

Figure 2 Release and metabolism of endogenous ATP and metabolism of exogenous e-ATP by the guinea-pig vas deferens in nominally Ca^{2+} -free bathing solution. Conditions were as for Fig. 1, except that Ca^{2+} was omitted from the bathing solution. **a**, Stimulation of the sympathetic nerves (EFS; 8 Hz for 60 s, indicated by thin horizontal bar) evoked only a small increase in adenosine overflow ($n = 4$). **b**, The metabolism of e-ATP was similar to that obtained in the presence of Ca^{2+} (compare **b** to Fig. 1c, d; $n = 4$). **c**, Sympathetic nerve stimulation in nominally Ca^{2+} -free bathing solution largely prevented the metabolism of e-ATP ($n = 4$; compare **c** with Fig. 1e, f).

e-purines during stimulation, their sum total was constant throughout (Fig. 1e, f). Therefore, there is no significant uptake of ATP or its metabolites and no further degradation of adenosine by the tissue. Electrical field stimulation in the absence of the vas deferens did not degrade e-ATP (not shown). Thus, sympathetic nerve stimulation greatly increases the extracellular breakdown of ATP.

One explanation for these results is that stimulation releases nucleotidases together with the neurotransmitters, in which case their release should be abolished under conditions that inhibit vesicular release. Confirming our previous report¹⁴, the nerve-evoked overflow of endogenous purines was virtually abolished when the experiments shown in Fig. 1a, b were repeated in a nominally Ca^{2+} -free solution (Fig. 2a). Likewise, the nerve-stimulation-evoked breakdown of exogenous e-ATP was attenuated, so that almost all of the e-purine was recovered as e-ATP (compare Fig. 1e, f with Fig. 2c). Results were similar when Cd^{2+} (300 μM) was added to the superfusate to block voltage-dependent calcium ion channels (not shown). These results are consistent with vesicular release of nucleotidases from sympathetic nerves.

To test whether the nucleotidases released by sympathetic nerve stimulation are soluble and stable in physiological solution, in all subsequent experiments we collected superfusate during sympathetic nerve stimulation and analysed its ability to metabolize exogenous e-ATP in the absence of tissue. This also has the advantage of eliminating any contribution by tissue ectonucleotidase to ATP degradation.

Superfusate collected from unstimulated tissues had no effect on exogenous e-ATP (Fig. 3a), but superfusate collected during nerve stimulation degraded exogenous e-ATP, as shown by the appearance of e-adenosine and, to a lesser extent, e-ADP and e-AMP (Fig. 3b). This effect was abolished when neurotransmission was inhibited by omitting Ca^{2+} from (Fig. 3c) or adding tetrodotoxin (1 μM) to (Fig.

3d) the solution bathing the tissues during nerve stimulation. Thus, sympathetic nerve stimulation releases nucleotidases into the extracellular space that can subsequently metabolize purines in the absence of tissue.

One possibility consistent with these results is that the nucleotidases are released from smooth-muscle cells during nerve stimulation when the neurotransmitters ATP and noradrenaline initiate contraction. However, the P2-receptor agonists ATP and α, β -methylene ATP and the α_1 -adrenoceptor agonists noradrenaline, methoxamine and phenylephrine (all at 1, 10, 100 μM) each elicited contraction, but none released nucleotidases into the bathing solution (not shown). This suggests that the source of the enzymes is neuronal.

An alternative possibility is that the nerves release a factor that cleaves a tether linking ectonucleotidases to the extracellular membrane of the smooth muscle cells^{15,16}, thereby activating them. We therefore did similar experiments in the absence of smooth muscle using cultured PC12 cells. Like postganglionic sympathetic neurones, these cells arise from the neural crest and display vesicular release of ATP and catecholamines. Depolarization of PC12 cells with KCl also evoked the release of ATP-metabolizing enzymes into the superfusate (S.M.-T. *et al.*, manuscript in preparation).

The chemical properties of the releasable nucleotidases are indicated in Fig. 4. For example, acidification of the incubation medium to pH 4 with perchloroacetic acid prevented the superfusate from breaking down e-ATP (100 nM (Fig. 4b). Enzyme activity was also abolished by heating the samples at 80 °C for 10 min and greatly reduced by 10-fold dilution (not shown).

The most important step in the termination of the neurotransmitter actions of ATP is the initial hydrolysis to produce ADP, which is much less active at P2X-receptors. To characterize the soluble ATPase, we investigated the effects of compounds that inhibit other

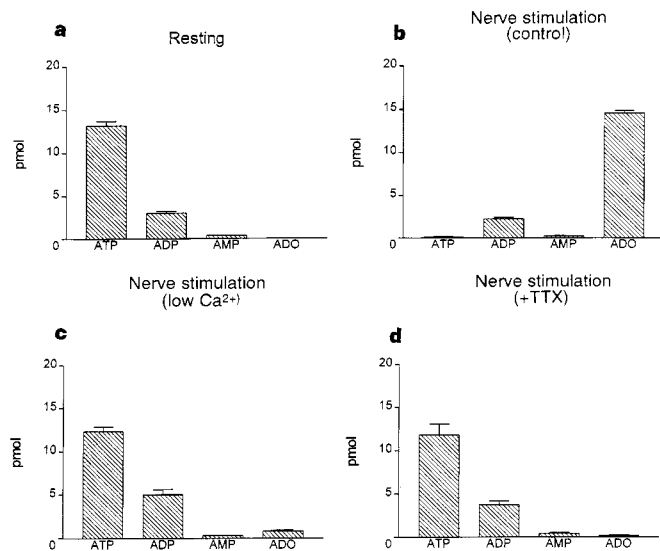


Figure 3 Release of soluble nucleotidases and the effect of nominally Ca^{2+} -free solution and tetrodotoxin. Samples of superfusate were collected for 20 s from resting tissues (no nerve stimulation) (a), or for the first 20 s of sympathetic nerve stimulation at 8 Hz (b–d). 100 μl of a 200 nM solution of e-ATP was then added to a 100- μl aliquot of the superfusate and e-ATP breakdown was monitored. Each panel shows the mean data for the amounts of individual e-purines present at the end of the incubation. a, Superfusate collected under resting conditions had minimal effect on e-ATP ($n = 4$); but in b, e-ATP was substantially metabolized by superfusate collected during nerve stimulation ($n = 6$). Adenosine was the main metabolite detected. This breakdown of e-ATP was inhibited by c, omission of Ca^{2+} ions ($n = 4$), or d, addition of tetrodotoxin (TTX; 1 μM ; $n = 4$) to the bathing solution during nerve stimulation.

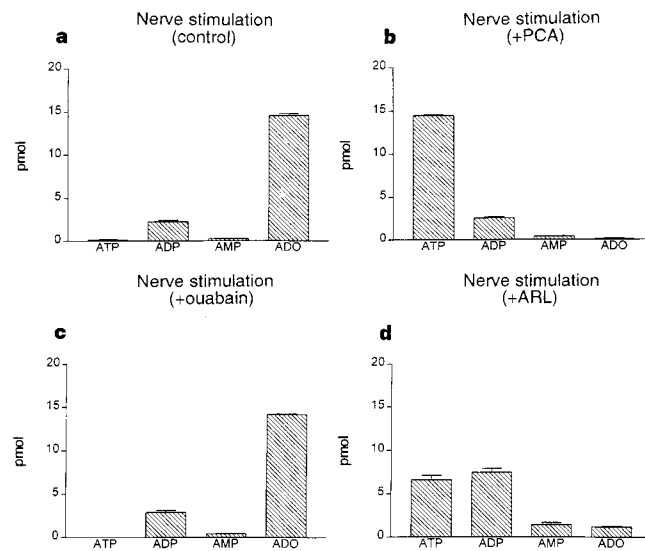


Figure 4 Inhibition of the soluble nucleotidase activity. Samples of superfusate were collected during the first 20 s of sympathetic nerve stimulation at 8 Hz and the breakdown of 100 nM e-ATP monitored in the absence (a) or presence (b–d) of inhibitors of ATPase enzymes. a, In the absence of inhibitors, e-ATP was largely broken down to e-adenosine. b, This effect was abolished by acidification to pH 4 by perchloroacetic acid (PCA); but in c, was unaffected by ouabain (10 μM); and in d, was partially inhibited by ARL 67156 (8-*N,N*-diethyl- β , γ -dibromomethyleneATP (from Astra Charnwood) ($n = 4$ each).

ATPases. Ouabain (10 μ M) an inhibitor of Na^+/K^+ -ATPase (EC 3.6.1.37) had no effect on e-ATP breakdown (Fig. 4c). Likewise, no effect was seen with orthovanadate (500 μ M), sodium azide (1 mM), oligomycin (5 mg ml⁻¹), *N*-ethyl maleimide (0.1–1 mM) and bafilomycin (100 nM) (not shown), which inhibit Na^+/K^+ -ATPase, Ca^{2+} -transport ATPase (EC 3.6.1.38) and mitochondrial ATPase (EC 3.6.1.34) but not ecto-ATPase (EC 3.6.1.3) (refs 3,4).

ARL 67156 inhibits ecto-ATPase in human blood cells¹⁷ and rat vas deferens¹⁸ and potentiates the purinergic component of neurogenic contractions in the guinea-pig vas deferens⁷. In the presence of ARL 67156 (100 μ M), almost half of the e-ATP added to superfusate samples was recovered intact (Fig. 4d). A similar amount was recovered as e-ADP, with much smaller amounts of e-AMP and e-adenosine. Breakdown of e-ATP was also depressed by non-derivatized ATP (100 μ M) and suramin (100 μ M), a P₂-receptor antagonist and inhibitor of ecto-ATPase^{18–21} (not shown). Thus, the properties of the releasable ATPase are similar to those of ecto-ATPase.

These studies show that when stimulated, sympathetic nerves release not only ATP and noradrenaline as neurotransmitters, but also soluble nucleotidases that can act in conjunction with ectonucleotidase to inactivate ATP. Thus, unlike other neurotransmitters, inactivation of ATP involves the neuronal release of specific metabolic enzymes. Enzyme release was Ca^{2+} -dependent, implying vesicular storage. This raises the question as to why vesicular ATP isn't inactivated before it is released. It could be that ATP and the enzymes are stored in separate vesicles, as proposed for ATP and noradrenaline¹⁴. Alternatively, ATP and the nucleotidases may be stored together in the same vesicles, but with the enzymes in an inactive state, perhaps due to the acidic pH of the vesicle interior²². Release into the less acidic extracellular space could then allow enzyme activation. The identities of the individual releasable enzymes are not known. The initial dephosphorylation of ATP was unaffected by inhibitors of ion-pump-associated ATPases, but was depressed by the ecto-ATPase inhibitors^{17–21} suramin and ARL 67156. Thus, the releasable ATPase may be structurally related to ecto-ATPase, although we cannot yet differentiate between ecto-ATPase and apyrase (which removes two phosphate groups from ATP). Several ectonucleotidases have now been cloned^{15,16,23–26} and these should be compared with the releasable enzymes. □

Methods

Tissue preparation. Male albino guinea-pigs (350–400 g) were killed by decapitation. The vasa deferentia were removed, cleaned of connective tissue and the lumen exposed by section along the longitudinal axis. Three tissues were loaded in a Brandel superfusion chamber (fluid volume, 200 μ l). Whatman 541 filters were cut to fit both ends of the chamber, which was then inserted vertically into a thermostatic block with two platinum 'screen' electrodes at either end¹⁴. The tissues were superfused from bottom to top at 2 ml min⁻¹ with Krebs solution (37 °C) of the following composition (in mM): NaCl 110, NaHCO₃ 24.8, KCl 4.6, MgSO₄ 1.2, KH₂PO₄ 1.2, CaCl₂ 2.5 and glucose 5.6, bubbled with 95% O₂, 5% CO₂, and allowed to equilibrate for 45 min. The role of Ca^{2+} ions was studied in some experiments by omitting CaCl₂ from the bathing solution. (Note that we use the term 'bathing solution' when referring to solution applied to the tissue and 'superfusate' when referring to solution collected once it has been in contact with the tissue.) Sympathetic nerves were stimulated by electrical field stimulation (EFS) at 8 Hz with a pulse width of 0.1 ms and supramaximal voltage.

HPLC assay. The adenine nucleotide and adenosine content of superfusate samples were analysed as described^{14,27}. Briefly, to measure the overflow of endogenous purines, samples of the superfusate (~320 μ l) were collected in ice-cold test-tubes at 10-s intervals, before, during and after sympathetic nerve stimulation. A control sample of the Krebs solution, which had not been in contact with the tissues, was also prepared. Chloroacetaldehyde was then added to form fluorescent 1^N-etheno derivatives (e-purines) of ATP, ADP, AMP and adenosine present in the samples. The e-purines were then separated on a gradient HPLC system equipped with a Waters Resolve radial pack cartridge (8Nv PH 4 μ ; 8 $\times$ 10 mm), and the amount of each was quantified using a

Shimadzu (RF 555) fluorescent monitor at an excitation wavelength of 230 nm and emission wavelength of 420 nm. Buffer solutions consisted of 0.1 M phosphate, pH 6.0 (buffer A) and 75% 0.1 M phosphate, pH 6.0, and 25% methanol (buffer B). The nucleosides and adenosine were separated using a gradient in which the concentration of buffer B was increased from 0 to 100% in 8 min. Identification of individual peaks was by comparison with the retention times of known e-purine standards and the concentration was determined by the peak area per pmol compared with standards. Standards were run with each set of samples. Results were normalized for volume and tissue weight and the data calculated as pmol mg⁻¹. A small amount of ADP (<10%) was present in freshly prepared ATP solutions that had not been in contact with tissues.

To study the metabolism of exogenous ATP when applied to the vas deferens, tissues were superfused with a solution containing 100 nM e-ATP. Samples of the superfusate were then collected at 10-s intervals and assayed as described. To investigate the activity of releasable soluble nucleotidases, samples of the superfusate were collected for 20 s under resting conditions, or for the first 20 s during sympathetic nerve stimulation. 100 μ l of a 200 nM solution of e-ATP were then added to a 100- μ l aliquot of the superfusate and the mixture assayed by HPLC. Results were calculated as pmol per sample.

Received 16 December 1996; accepted 6 March 1997.

- Amara, S. G. & Kuhar, M. J. Neurotransmitters: recent progress. *Annu. Rev. Neurosci.* **16**, 73–93 (1993).
- Bowman, W. C. & Rand, M. J. *Textbook of Pharmacology* 2nd edn., 10.1–10.8 (Blackwell, Oxford, 1980).
- Ziganshin, A. U., Hoyle, C. H. V. & Burnstock, G. Ecto-enzymes and metabolism of extracellular ATP. *Drug Dev. Res.* **32**, 134–146 (1994).
- Plesner, L. Ecto-ATPases: identities and functions. *Int. J. Cytol.* **158**, 141–214 (1995).
- Edwards, F. A., Gibb, A. J. & Colquhoun, D. ATP receptor-mediated synaptic currents in the central nervous system. *Nature* **359**, 144–146 (1992).
- Burnstock, G. Purinergic mechanisms. *Annals NY Acad. Sci.* **603**, 1–17 (1990).
- Westfall, T. D., Kennedy, C. & Sneddon, P. Enhancement of sympathetic purinergic neurotransmission in the guinea-pig isolated vas deferens by the novel ecto ATPase inhibitor ARL 67156. *Br. J. Pharmacol.* **117**, 867–872 (1996).
- Sneddon, P., McLaren, G. J. & Kennedy, C. Purinergic cotransmission: sympathetic nerves. *Semin. Neurosci.* **8**, 201–205 (1996).
- Gordon, J. Extracellular ATP: effects, sources and fate. *Biochem. J.* **233**, 309–319 (1986).
- Kennedy, C. & Leff, P. How should P₂-purinoceptors be characterized pharmacologically? *Trends Pharmacol. Sci.* **16**, 168–174 (1995).
- Pearson, J. D., Coade, S. B. & Cusack, N. J. Characterization of ectonucleotidases on vascular smooth muscle cells. *Biochem. J.* **230**, 503–507 (1985).
- Welford, L. A., Cusack, N. J. & Hourani, S. M. O. ATP analogues and the guinea-pig taenia coli: a comparison of the structure activity relationships of ectonucleotidases with those of the P₂-purinoceptor. *Eur. J. Pharmacol.* **129**, 217–224 (1986).
- Welford, L. A., Cusack, N. J. & Hourani, S. M. O. The structure-activity relationships of ectonucleotidases and of excitatory P₂-purinoceptors: evidence that dephosphorylation of ATP analogues reduces pharmacological potency. *Eur. J. Pharmacol.* **141**, 123–130 (1987).
- Todorov, L. D., Mihaylova-Todorova, S., Craviso, G. L., Bjur, R. A. & Westfall, D. P. Evidence for the differential release of the cotransmitters ATP and noradrenaline from sympathetic nerves of the guinea-pig vas deferens. *J. Physiol. (Lond)* **496**, 731–748 (1996).
- Misumi, Y., Ogata, S., Hirose, S. & Ikehara, Y. Primary structure of rat liver 5' nucleotidase deduced from the cDNA. *J. Biol. Chem.* **265**, 2178–2183 (1990).
- Misumi, Y., Ogata, S., Ohkubo, K., Hirose, S. & Ikehara, Y. Primary structure of human placental 5' nucleotidase and identification of the glycolipid anchor in the mature form. *Eur. J. Biochem.* **191**, 563–569 (1990).
- Crack, B. E. *et al.* Pharmacological and biochemical analysis of FPL 67156, a novel, selective inhibitor of ecto-ATPase. *Br. J. Pharmacol.* **114**, 475–481 (1995).
- Khakh, B. S., Michel, A. D. & Humphrey, P. P. A. Inhibition of ectoATPase and Ca-ATPase in rat vas deferens by P₂-purinoceptor antagonists. *Br. J. Pharmacol.* **115**, 2P (1995).
- Hourani, S. M. O. & Chown, J. A. The effects of some possible inhibitors of ectonucleotidases on the breakdown and pharmacological effects of ATP in the guinea-pig urinary bladder. *Gen. Pharmacol.* **20**, 413–416 (1989).
- Bailey, S. J. & Hourani, S. M. O. Differential effects of suramin on P₂-purinoceptors mediating contraction of the guinea-pig vas deferens and urinary bladder. *Br. J. Pharmacol.* **112**, 219–225 (1994).
- Stout, J. G. & Kirley, T. L. Inhibition of purified chicken gizzard smooth muscle ecto-ATPase by P₂-purinoceptor antagonists. *Biochem. Mol. Biol. Int.* **36**, 927–934 (1995).
- Schuldiner, S., Shirvan, A. & Linal, M. Vesicular neurotransmitter transporters: from bacteria to humans. *Physiol. Rev.* **95**, 369–392 (1995).
- Maliszewski, C. R. *et al.* The CD39 lymphoid cell activation antigen. Molecular cloning and structural characterization. *J. Immunol.* **153**, 3574–3583 (1994).
- Bermudes, D., Ran Peck, K., Afifi Afifi, M., Beckers, C. J. M. & Joiner, K. A. Tandemly repeated genes encode nucleoside triphosphate hydrolase isoforms secreted into the parasitophorous vacuole of *Toxoplasma gondii*. *J. Biol. Chem.* **269**, 29252–29260 (1994).
- Asai, T., Miura, S., Sibley, L. D., Okabayashi, H. & Tkeuchi, T. Biochemical and molecular characterization of nucleoside triphosphate hydrolase isozymes from the parasitic protozoan *Toxoplasma gondii*. *J. Biol. Chem.* **270**, 11391–11397 (1995).
- Wang, T. F. & Guidotti, G. CD39 is an ecto-(Ca²⁺ Mg²⁺)-ATPase. *J. Biol. Chem.* **271**, 9898–9901 (1996).
- Levitt, B., Head, R. J. & Westfall, D. P. High-pressure liquid chromatographic-fluorometric detection of adenosine and adenine nucleotides: applications to endogenous and electrically induced release of adenyli purines in guinea-pig vas deferens. *Analyt. Biochem.* **137**, 93–100 (1984).

Acknowledgements. This work was supported by grants from the NIH and the American Heart Association (D.P.W., L.D.T., S.M.-T.), by Astra plc and the Wellcome Trust (C.K., P.S., T.D.W.).

Correspondence and requests for materials should be addressed to D.P.W. (e-mail: westfall@admin.unn.edu).

Connexin 26 mutations in hereditary non-syndromic sensorineural deafness

D. P. Kelsell[†], J. Dunlop^{*}, H. P. Stevens^{*}, N. J. Lench[‡], J. N. Liang[§], G. Parry^{||}, R. F. Mueller[‡] & I. M. Leigh^{*}

^{*} Academic Department of Dermatology, St Bartholomew's and the Royal London School of Medicine and Dentistry, Queen Mary and Westfield College, 2 Newark Street, London E1 2AT, UK

[‡] Molecular Medicine Unit, St James's University Hospital, Leeds LS9 7TF, UK

[§] The Temporal Bone Laboratory, The Institute of Laryngology and Otolaryngology, University College London, 330/332 Gray's Inn Road, London, WC1X 8EE, UK

^{||} Department of Paediatrics, St Luke's Hospital, Bradford, BD5 0NA, UK

Severe deafness or hearing impairment is the most prevalent inherited sensory disorder, affecting about 1 in 1,000 children¹. Most deafness results from peripheral auditory defects that occur as a consequence of either conductive (outer or middle ear) or sensorineural (cochlea) abnormalities. Although a number of mutant genes have been identified that are responsible for syndromic (multiple phenotypic disease) deafness such as Waardenburg syndrome and Usher 1B syndrome²⁻⁴, little is known about the genetic basis of non-syndromic (single phenotypic disease) deafness. Here we study a pedigree containing cases of autosomal dominant deafness and have identified a mutation in the gene encoding the gap-junction protein connexin 26 (Cx26) that segregates with the profound deafness in the family. Cx26 mutations resulting in premature stop codons were also found in three autosomal recessive non-syndromic sensorineural deafness pedigrees, genetically linked to chromosome 13q11-12 (DFNB1), where the Cx26 gene is localized. Immunohistochemical staining of human cochlear cells for Cx26 demonstrated high levels of expression. To our knowledge, this is the first non-syndromic sensorineural autosomal deafness susceptibility gene to be identified, which implicates Cx26 as an important component of the human cochlea.

Initially we studied a caucasian pedigree in which both autosomal dominant palmoplantar keratoderma (PPK, in which there is abnormal callusing of palms and soles) and congenital sensorineural deafness were segregating^{5,6}. From a clinical reevaluation, it is unlikely that a mutation in the same gene results in the PPK and the profound deafness phenotype in this pedigree as there are three individuals who have the PPK but not the profound hearing disorder (Fig. 1). Haplotype analysis suggested linkage with the skin disease phenotype and microsatellite DNA marker loci mapping to 13q11-12 (data not shown). Three disease loci have been localized to 13q11-12; non-syndromic deafness (autosomal recessive DFNB1 and autosomal dominant DFNA3; refs 7,8) and the autosomal dominant skin disease Clouston hidrotic ectodermal dysplasia (HED⁹).

A candidate gene for both the Clouston's HED and the PPK phenotypes was that encoding connexin 26 (GJB2), a β -type connexin, which maps to 13q11-12 (ref. 10). The connexins are a large family of proteins with four transmembrane domains, which have been implicated in gap-junctional intercellular communication^{11,12}. Connexins are widely expressed in vertebrates, with distinct but overlapping expression patterns for each member of the family¹³. Immunohistological studies in the mouse have shown that Cx26 is expressed in the suprabasal layer of the epidermis^{14,15}. Therefore, on the basis of its chromosomal localization and its expression pattern, we decided to look for Cx26

mutations in affected individuals from two pedigrees with Clouston's HED and the pedigree with PPK and deafness. The single coding exon of Cx26 was amplified by the polymerase chain reaction (PCR) and the resulting products sequenced.

No sequence variants were detected in affected members of the Clouston's HED pedigrees. However, a T $\rightarrow$ C substitution in codon 34 (exon 1) was identified in two affected individuals from the PPK-and-deafness pedigree (Fig. 2a). This substitution resulted in a methionine (ATG)-to-threonine (ACG) (M34T) change in the translated sequence and was not observed in 80 unrelated chromosomes. This methionine is located in the first putative transmembrane domain of Cx26 and is conserved between different Cx26 and connexin 32 species, whereas the other β -connexins contain a leucine at position 44, a conservative amino-acid substitution¹⁶. This suggests that the Met at codon 34 is important for the activity of Cx26. The M34T mutation results in a bulky hydrophobic residue being replaced by a polar residue and is likely to alter the first transmembrane domain. This mutation may act in a dominant-negative fashion. However, the M34T mutation segregated with the profound-deafness phenotype but not the skin disorder in this family (Fig. 1b). These data indicated that the M34T mutation was the genetic basis of the autosomal dominant deafness but not the PPK in this pedigree.

To investigate the hypothesis that Cx26 mutations can give rise to deafness, we examined the expression pattern of Cx26 using murine monoclonal antibodies on sections of normal human interfollicular and palmoplantar epidermis, psoriasis and the cochlea. In contrast to that found in the mouse, Cx26 staining was only weak in human skin but expression was stronger in the psoriatic epidermis (data not shown). However, Cx26 expression was demonstrated in the stria vascularis, basement membrane, limbus and the spiral prominence of the cochlea (Fig. 3a, c). Identification of the human cochlea was achieved by examination of serial haematoxylin-and-eosin-stained frozen sections. Unfortunately the hair cells and part of the Reissner's membrane were lost during processing. However, fragments of the Reissner's membrane can still be seen and the basement membrane on which the hair cells lie has been clearly demonstrated (Fig. 3b). The stria vascularis was also identified, as was the spiral prominence, limbus and cochlear nerve. To confirm that human cochlear duct had been obtained, we looked for the presence of simple epithelia that lines the duct. Staining was achieved using mouse monoclonal antibodies specific for the simple epithelial keratins (K7, K8, K18 and K19) and the suprabasal keratins (K1 and K10; ref. 17). Staining for the keratins K1 and K10 was negative. The positive staining pattern for K7, K8, K18 and K19 within the stria vascularis confirmed the presence of a simple epithelial-derived structure (data not shown). Cx26 expression has been reported in the rat cochlea, particularly in the cells lining the cochlear duct and the stria vascularis¹⁸. The distribution of Cx26 staining with the cochlea of the rat and that of the human appears to be similar. These immunohistochemical data are consistent with the idea that the Cx26 mutation is the underlying cause of deafness in the pedigree and suggest the Cx26 gene is the non-syndromic autosomal dominant sensorineural deafness locus, DFNA3, but the identification of further Cx26 mutations in DFNA3-linked deafness pedigrees are required to confirm this.

As the recessive deafness locus DFNB1 also maps to the same region on chromosome 13 as Cx26 (ref. 7), we investigated the possibility that Cx26 mutations could cause deafness in affected individuals from DFNB1-linked pedigrees. A gene in a large consanguineous family of Pakistani origin with recessive non-syndromic profound deafness had been localized to 13q11-12 (DFNB1; ref. 19). None of the family members had any cutaneous abnormalities. DNA from two affected individuals and their parents were screened for mutations in Cx26. Both affected individuals were homozygous for a G-to-A substitution in codon 77 which results in a premature stop codon (Trp to Opal: W77X) (Fig. 2c). The W77X

[†] Present address: Biopharmaceuticals, Smithkline Beecham, New Frontiers Science Park, Third Avenue, Harlow, Essex CM19 5AW, UK.

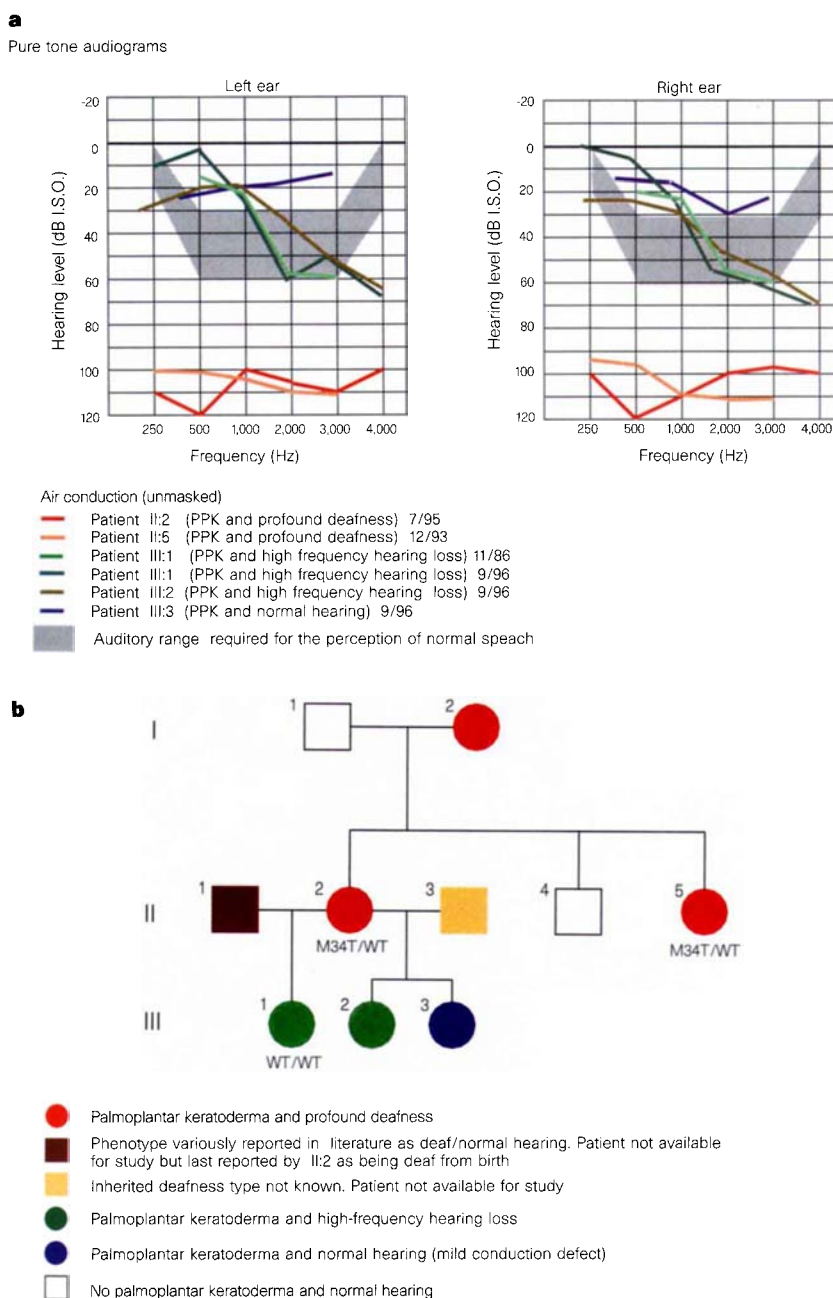


Figure 1 The phenotypes in the autosomal dominant deafness pedigree associated with palmoplantar keratoderma has been changed from that previously published^{5,6}. This change in phenotype was achieved by patient interview, audiographic examination and retrieval of historic audiograms. The fathers (II:1 and II:3 in **b**), although unavailable for interview and audiographic examination, were reported by II:2 as having been deaf from birth and communicated by sign language. **a**, Audiograms of assessed individuals. This demonstrates that two distinct types of deafness segregate within this family. The first is a profound deafness giving a flat trace on audiographic examination (individuals II:2, aged 36 years, and II:5, aged 24 at date tested, both of whom were educated in deaf schools and retain no speech, suggesting that the deafness was not progressive). The second is a high-frequency hearing loss (individuals III:1 and III:2) in two of the children of II:2, both of whom speak normally, can communicate over the telephone and are being educated in normal schools. In individual III:1 the hearing loss has been stable over a 10-year period (audiogram traces in

November 1986 at age 3 years and in September 1996 at 13 years) making the suggestion unlikely that the high-frequency hearing loss may progress to profound deafness. Individual III:3 (aged 4 years) has previously been recorded as having both PPK and deafness, but the audiogram shows clearly that she has only a mild conduction defect with no evidence of either the profound deafness seen in the mother (II:2) or the high-frequency hearing loss seen in her two sisters (III:1 and III:2, aged 6 years). The date when the air conduction audiograms were performed is shown in months and years after the patient details in the key for **a**. Ages are given at the date of audiograms. **b**, Structure of the autosomal dominant deafness pedigree in which PPK is also segregating. The M34T mutation in connexin 26 was only present (for the family members for which a DNA sample was available) in the two female individuals with both profound deafness and PPK. WT, wild-type allele. Squares indicate males and circles females. Unfilled symbols indicate neither PPK nor deafness.

mutation is in the second transmembrane domain of Cx26. The parents were heterozygous for this mutation and had no noticeable hearing impairment¹⁹. This sequence variant was not present in 80 unrelated chromosomes.

Additional evidence that Cx26 is the DFNB1 gene came from the analysis of two further consanguineous Pakistani families with non-syndromic profound deafness. Disease in these pedigrees showed supportive evidence of linkage to 13q11–12 (data not shown). Two affected individuals from each pedigree were screened for Cx26 mutations. All four individuals were homozygous for a G → A

substitution at codon 24 that results in a premature stop codon (Trp to Amber: W24X) and was not present in 80 unrelated chromosomes (Fig. 2b). The W24X mutation occurs in the first transmembrane domain of CX26. Disease haplotype comparison indicates that these two identical mutations arose independently of each other (data not shown).

Genetic linkage studies have identified at least eighteen regions of the human genome that harbour a susceptibility locus for non-syndromic sensorineural deafness²⁰. So far, none of these disease genes has been identified. We have described both autosomal

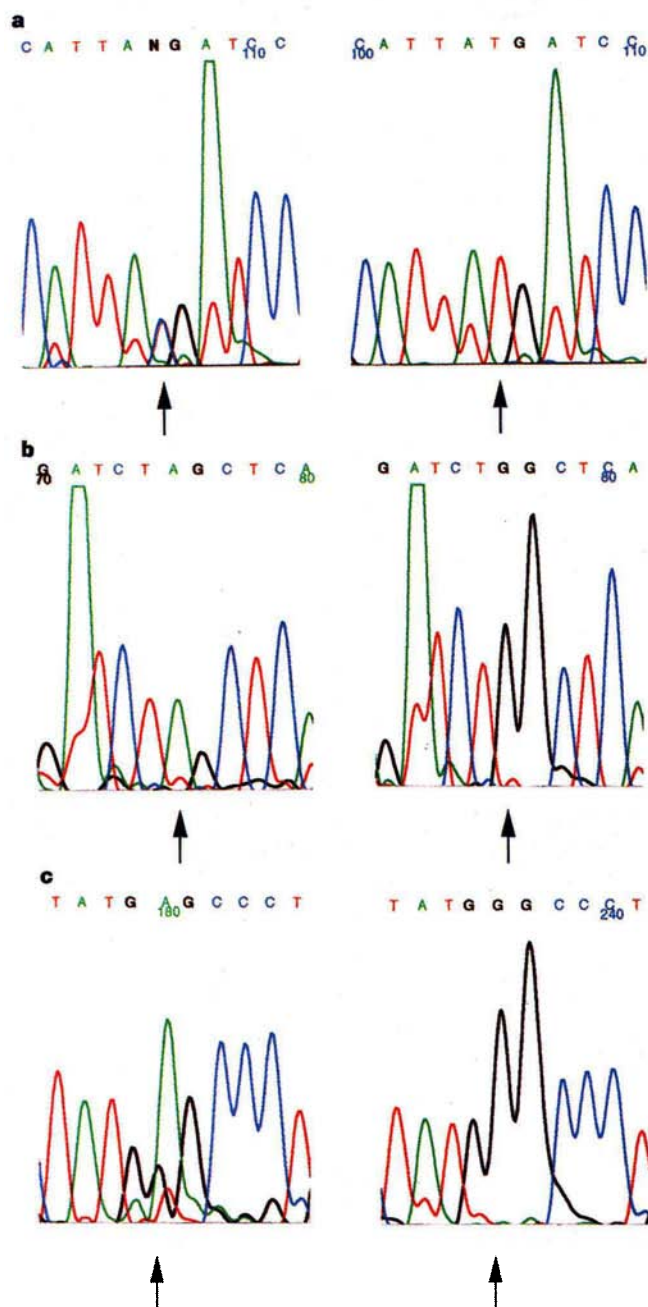


Figure 2 Mutation analysis of Cx26. Representative sequence electropherograms of the three distinct germline mutations identified: **a**, M34T in the autosomal dominant sensorineural deafness pedigree; **b**, W24X; and **c**, W77X in the autosomal recessive sensorineural deafness pedigrees. Corresponding portions of Cx26 genomic DNA sequences from left, an affected individual, and right, a non-carrier. The position of each mutation is indicated by an arrow. N denotes an unknown base (in this case, heterozygous for a C and T nucleotide) as determined by the sequence analysis package.

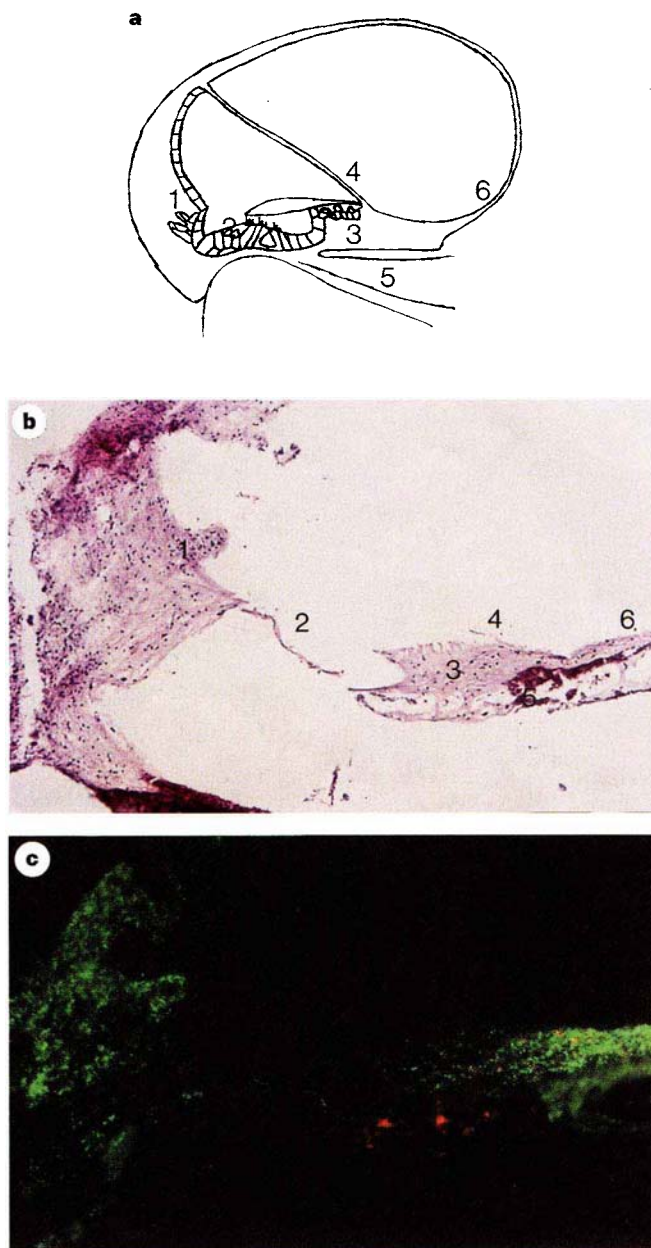


Figure 3 Cx26 immunofluorescence staining of tissue derived from the cochlea duct. **a**, Line diagram depicting the cochlea: 1, spiral prominence; 2, organ of corti; 3, limbus; 4, Reissner's membrane; 5, cochlear nerve; 6, stria vascularis. **b**, Haematoxylin and eosin staining of the same region (magnification, 10×). **c**, Rabbit polyclonal anti-Cx26 staining showing staining of the stria vascularis, basement membrane, limbus, spiral prominence of the cochlea. The hair cells and most of the Reissner's membrane were lost in the processing of the sample (magnification, 16×).

dominant and recessive Cx26 mutations in pedigrees with non-syndromic deafness. In addition, Cx26 expression was demonstrated in human cochlear cells. Therefore, combined mutation and expression data implicate Cx26 as the gene responsible for deafness in these and possibly other families linked to 13q11–12 (DFNA3 and DFNB1), making it the first non-syndromic sensorineural deafness gene to be discovered. A role for gap junctions in the recycling of potassium ions back to the endolymph of the cochlear duct after stimulation of the sensory hair cells has been suggested¹⁸, and the loss of Cx26 would be expected to disrupt this potassium ion flow, leading to the loss of hearing. A feature of the individuals homozygous for Cx26 truncations was the lack of any other observable phenotype, even though Cx26 is widely expressed in other tissues, for example, the liver and the pancreas¹¹. It is unclear how the M34T mutation leads to the dominant-negative phenotype, or if functional channels can form and how these might interact with wild-type channels. This is the third human disease in which connexin mutations have been reported, the others being Cx32 in X-linked Charcot-Marie-Tooth disease (which may also be associated with deafness)²¹ and Cx43 in viscerotaxial heterotaxia syndrome²². The identification of connexin mutations in these and other human diseases will aid our understanding of their biology. □

Methods

Haplotype and linkage analysis. Lymphocyte DNA extracted from family members was genotyped by PCR amplification of microsatellite marker loci mapping to 13q11–12 with fluorescently labelled primers. Primer sequences were obtained from the Genome Database. The PCR products were run on a 373A Automated DNA sequencer running Genescan software (Applied Biosystems). Alleles were analysed using Genotyper software (Applied Biosystems). Linkage analysis has been described¹⁹.

Screening of Cx26 gene for mutations. The coding exon was amplified using PCR and specific primers that produced two overlapping PCR products. The four primer sequences used were: 5'-TCTTTTCCAGAGCAACCGC-3' with 5'-GACACGAAGATCAGCTGCAG-3' and 5'-GGGCAATGCGTTAACTGGC-3' with 5'-CCAGGCTGCAAGAACGTGTG-3'. The Cx26 sequence was taken from GenBank accession number M86849. PCR conditions were 35 cycles of denaturation at 95°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for 30 s. PCR products were band-purified on a 1% low-melting-point agarose gel using Wizard spin purification columns (Promega). The products were sequenced directly from both ends using an AmpliTaq Cycle sequencing kit and ABI Prism 377 DNA sequencer.

Immunohistochemical analysis. Skin biopsy material was obtained from the non-dominant hypothenar eminence, abdomen and psoriatic plaques from volunteers and snap-frozen in liquid nitrogen. Human cochlea was obtained post mortem from a 76-year-old woman 36 h after death. The cochlea was identified by serial microdissection of the temporal bone and was removed by microdissection. The cochlea was embedded in Cryo-M-Block (Bright) and snap-frozen in liquid nitrogen. Immunohistochemical analysis was performed on serial 5-µm cryostat sections using a standard technique¹⁷. Every fifth section was stained with haematoxylin and eosin to assess orientation. The polyclonal rabbit anti-connexin 26 antibody (Zymed) was used at a 1/100 dilution.

Received 14 January; accepted 20 March 1997.

1. Morton, N. E. Genetic epidemiology of hearing impairment. *Ann. N.Y. Acad. Sci.* **630**, 16–31 (1991).
2. Tassabehji, M. et al. Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355**, 635–636 (1992).
3. Tassabehji, M. et al. Mutations in the PAX3 gene causing Waardenburg syndrome type 1 and type 2. *Nature Genet.* **3**, 26–30 (1993).
4. Weil, D. et al. Defective myosin VIIA gene responsible for Usher syndrome 1B. *Nature* **374**, 60–61 (1995).
5. Verbov, J. Palmoplantar keratoderma, deafness and atopy. *Br. J. Dermatol.* **116**, 881–882 (1987).
6. Fitzgerald, D. A. & Verbov, J. L. Hereditary palmoplantar keratoderma with deafness. *Br. J. Dermatol.* **134**, 939–942 (1996).
7. Guilford, P. et al. A human gene responsible for neurosensory, non-syndromic recessive deafness is a candidate homologue of the mouse *sh-1* gene. *Hum. Mol. Genet.* **3**, 989–993 (1994).
8. Chaib, H. et al. A gene responsible for a dominant form of neurosensory non-syndromic deafness maps to NSRD1 recessive deafness gene interval. *Hum. Mol. Genet.* **3**, 2219–2222 (1994).
9. Kibar, Z. et al. The gene responsible for Clouston hidrotic ectodermal dysplasia maps to the pericentromeric region of chromosome 13q. *Hum. Mol. Genet.* **5**, 543–547 (1996).
10. Mignon, C. et al. Assignment of connexin 26 (GJB2) and 46 (GJA3) genes to human chromosome

- 13q11–q12. *Cytogenet. Cell Genet.* **72**, 185–186 (1996).
11. Kumar, N. M. & Gilula, N. B. The gap junction communication channel. *Cell* **84**, 381–388 (1996).
12. Beyer, E. C. Gap junction. *Int. Rev. Cytol.* **137**, 231–242 (1993).
13. Kumar, N. M. & Gilula, N. B. Molecular biology and genetics of gap junction channels. *Semin. Cell Biol.* **3**, 3–16 (1992).
14. Rieck, B., Klier, F. G. & Gilula, N. B. Multiple gap junction genes are utilized during rat skin and hair development. *Development* **116**, 639–651 (1992).
15. Goliger, J. A. & Paul, D. L. Expression of gap junction proteins Cx26, Cx31.1, Cx37, and Cx43 in developing and mature epidermis. *Dev. Dyn.* **200**, 1–13 (1994).
16. Yeager, M. & Nicholson, B. J. Structure of gap junction intercellular channels. *Curr. Opin. Struct. Biol.* **6**, 183–192 (1996).
17. Markey, A. C., Lane, E. B., MacDonald, D. M. & Leigh, I. M. Keratin expression in basal cell carcinomas. *Br. J. Dermatol.* **126**, 154–160 (1992).
18. Kikuchi, T., Kimar, R. S., Paul, D. L. & Adams, J. C. Gap junctions in the rat cochlea: immunohistochemical and ultrastructural analysis. *Anat. Embryol.* **191**, 101–118 (1995).
19. Brown, K. A. et al. Linkage studies in non-syndromic recessive deafness (NSRD) in a family originating from the Mirpur region of Pakistan maps DFNB1 centromeric to D13S175. *Hum. Mol. Genet.* **5**, 169–173 (1996).
20. Petit, C. Genes responsible for human hereditary deafness: a symphony of a thousand. *Nature Genet.* **14**, 385–391 (1996).
21. Bergoffen, J. et al. Connexin mutations in X-linked Charcot-Marie-Tooth Disease. *Science* **262**, 2039–2041 (1993).
22. Britz-Cunningham, S. H., Shah, M. M., Zuppan, C. W. & Fletcher, W. H. Mutations of the connexin-43 gap junction gene in patients with heart malformations and defects of laterality. *N. Engl. J. Med.* **332**, 1323–1329 (1995).

Acknowledgements. We thank P. Spink and C. Chronnell for technical support; J. Verbov and R. E. Kelsell for DNA samples; D. J. Roberts, J. Verbov and H. Forsythe for audiograms; and J. Pitts, M. Finbow and I. A. McKay for discussion. This work was supported in part by grants from the Wellcome Trust (D.P.K., H.P.S., N.J.L., R.E.M. and I.M.L.), the MRC (J.D.) and the ICRF (H.P.S., I.M.L.).

Correspondence and requests for materials should be addressed to D.P.K. (e-mail: David_P_Kelsell@sphrd.com.uk).

Regulation of skeletal muscle mass in mice by a new TGF-β superfamily member

Alexandra C. McPherron*, Ann M. Lawler† & Se-Jin Lee*

* Department of Molecular Biology and Genetics, and † Department of Gynecology and Obstetrics, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA

The transforming growth factor-β (TGF-β) superfamily encompasses a large group of growth and differentiation factors playing important roles in regulating embryonic development and in maintaining tissue homeostasis in adult animals¹. Using degenerate polymerase chain reaction, we have identified a new murine TGF-β family member, growth/differentiation factor-8 (GDF-8), which is expressed specifically in developing and adult skeletal muscle. During early stages of embryogenesis, GDF-8 expression is restricted to the myotome compartment of developing somites. At later stages and in adult animals, GDF-8 is expressed in many different muscles throughout the body. To determine the biological function of GDF-8, we disrupted the GDF-8 gene by gene targeting in mice. GDF-8 null animals are significantly larger than wild-type animals and show a large and widespread increase in skeletal muscle mass. Individual muscles of mutant animals weigh 2–3 times more than those of wild-type animals, and the increase in mass appears to result from a combination of muscle cell hyperplasia and hypertrophy. These results suggest that GDF-8 functions specifically as a negative regulator of skeletal muscle growth.

To identify new members of the TGF-β superfamily, we designed degenerate oligonucleotides corresponding to conserved regions among the known family members and used these oligonucleotides as primers for polymerase chain reaction (PCR) on mouse genomic DNA. Among the new sequences obtained in this screen was one that we designated GDF-8. A full-length GDF-8 complementary DNA clone obtained from screening a murine skeletal muscle library contained a single long open reading frame encoding a protein 376 amino acids long. The predicted GDF-8 amino-acid sequence contains all of the hallmarks characteristic of members of

the TGF- β superfamily, including a signal sequence for secretion, a proteolytic processing site, and a carboxy-terminal region containing the conserved pattern of cysteine residues (Fig. 1a). Like the TGF- β s and inhibin- β s, GDF-8 contains nine cysteine residues in the C-terminal region. However, on the basis of pairwise comparisons of the C-terminal region of GDF-8 with those of the other members of the TGF- β superfamily, GDF-8 does not appear to fall into any of the known subfamilies of highly related factors, including the TGF- β s, inhibins and bone morphogenetic proteins (BMPs) (Fig. 1b). In the C-terminal region, GDF-8 is most homologous to Vgr-1 (45% amino-acid identity).

Like most other TGF- β family members, GDF-8 also appears to be highly conserved across species. By genomic Southern analysis, homologous sequences were detected in all mammalian species examined as well as in chickens (Fig. 1c). In most species, the GDF-8 probe also detected a second, more faintly hybridizing fragment. Further analysis of GDF-8-homologous sequences in murine and human DNA revealed that this second fragment corresponds to a highly related gene, GDF-11 (A.C.M. and S.-J.L., unpublished results); the isolation of GDF-11 and characterization of GDF-11 knock-out mice will be described elsewhere.

To determine whether the processing signals in the GDF-8 sequence are functional and whether GDF-8 forms dimers as do other members of the TGF- β superfamily, the GDF-8 cDNA was stably expressed in Chinese hamster ovary (CHO) cells. As shown in Fig. 1d, western analysis of conditioned medium prepared from these cells using an antiserum raised against a bacterially expressed C-terminal fragment of GDF-8 detected two protein species with apparent relative molecular masses (M_r) of about 52,000 (52K) and 15,000 (15K) under reducing conditions, consistent with unprocessed and processed forms of GDF-8, respectively. No bands were obtained either with preimmune serum or with conditioned medium from CHO cells transfected with an antisense construct (data not shown). Under non-reducing conditions, the GDF-8 antiserum detected two predominant protein species with apparent M_r s of about 101K and 25K, consistent with dimeric forms of unprocessed and processed GDF-8, respectively. Hence, like other TGF- β family members, GDF-8 appears to be secreted and proteolytically processed, and the C-terminal region appears to be capable of forming a disulphide-linked dimer.

To determine the expression pattern of GDF-8, *in situ* hybridization was performed on mouse embryos isolated at various stages of development. At all stages examined, the expression of GDF-8 messenger RNA appeared to be restricted to developing skeletal muscle. At early stages, GDF-8 expression was restricted to developing somites. By whole-mount *in situ* hybridization analysis, GDF-8 mRNA could first be detected as early as day 9.5 post-coitum (p.c.) in about one-third of the somites (Fig. 2a). At this stage of development, hybridization appeared to be restricted to the most mature (9 out of 21 in the example shown) rostral somites. By day 10.5 p.c. (Fig. 2b), GDF-8 expression was clearly evident in almost every somite (26 out of 33 in the example shown). On the basis of *in situ* hybridization analysis of sections prepared from day 10.5 p.c. embryos, the expression of GDF-8 in somites appeared to be localized to the myotome compartment (Fig. 2c). At later stages of development, GDF-8 expression was detected in a wide range of developing muscles. Examples of GDF-8 expression in developing muscles of the foot and hip region at day 14.5 p.c. are shown in Fig. 2d, e.

GDF-8 continues to be expressed in adult animals as well. By northern analysis, GDF-8 mRNA expression was seen almost exclusively in skeletal muscle among the different adult tissues examined (Fig. 2f). A significantly lower although clearly detectable signal was also seen in adipose tissue. On the basis of northern analysis of RNA prepared from a large number of different adult skeletal muscles, GDF-8 expression appeared to be widespread although the expression levels varied among individual muscles (Fig. 2g).

To determine the biological function of GDF-8, we disrupted the GDF-8 gene by homologous targeting in embryonic stem cells. To ensure that the resulting mice would be null for GDF-8 function, the entire mature C-terminal region was deleted and replaced by a neo cassette (Fig. 3a). Homologous targeting of the GDF-8 gene was seen in 13/131 gancyclovir/G418 doubly resistant embryonic stem (ES) cell clones. Following injection of these targeted clones into blastocysts, we obtained chimaeras from five independently derived ES clones that produced heterozygous pups when crossed to C57BL/6 females. Genotypic analysis (Fig. 3b) of 678 offspring derived from crosses of F1 heterozygotes showed 170+/+ (25%), 380+/- (56%) and 128-/- (19%). Although the ratio of genotypes was close to the expected ratio of 1:2:1, the smaller than expected number of homozygous mutants appeared to be statistically significant

Table 1 Total body weights of homozygous, heterozygous and wild-type animals

Genotype	Weight (g)			
	2 months	3 months	4 months	5 months
Male				
+/+	28.0 $\pm$ 3.0 (n = 18)	30.6 $\pm$ 3.7 (n = 35)	33.2 $\pm$ 4.4 (n = 26)	33.9 $\pm$ 3.9 (n = 20)
+/-	29.1 $\pm$ 2.9 (n = 53)	31.9 $\pm$ 4.0 (n = 90)	34.6 $\pm$ 3.9 (n = 68)	36.1 $\pm$ 3.3 (n = 59)
-/-	36.2 $\pm$ 4.0 (n = 23)*	41.0 $\pm$ 4.5 (n = 33)*	43.1 $\pm$ 3.5 (n = 23)*	42.9 $\pm$ 5.0 (n = 18)*
Female				
+/+	22.1 $\pm$ 2.0 (n = 12)	23.7 $\pm$ 3.2 (n = 35)	26.1 $\pm$ 3.6 (n = 16)	27.2 $\pm$ 3.3 (n = 19)
+/-	23.1 $\pm$ 2.6 (n = 31)	26.0 $\pm$ 2.9 (n = 53)*	27.6 $\pm$ 3.3 (n = 36)	28.8 $\pm$ 3.6 (n = 35)
-/-	26.2 $\pm$ 2.3 (n = 10)*	30.8 $\pm$ 3.8 (n = 21)*	32.7 $\pm$ 3.9 (n = 12)*	35.1 $\pm$ 3.8 (n = 15)*

* Significant difference between +/+ and -/- ($P < 0.001$).

† Significant difference between +/+ and +/- ($P < 0.01$).

Table 2 Weights of individual muscles from mutant and wild-type animals

	Weight (g)*		Per cent of +/-
	+/+ (n = 9)	-/- (n = 9)	
Body	30.3 $\pm$ 3.3	40.9 $\pm$ 4.2	135
Digastric	0.022 $\pm$ 0.005	0.045 $\pm$ 0.007	205
Pectoralis	0.178 $\pm$ 0.042	0.467 $\pm$ 0.067	262
Triceps brachii	0.158 $\pm$ 0.036	0.372 $\pm$ 0.039	235
Quadriceps	0.232 $\pm$ 0.052	0.470 $\pm$ 0.053	203
Gastrocnemius/plantar	0.150 $\pm$ 0.033	0.328 $\pm$ 0.020	219
Tibialis cranialis	0.047 $\pm$ 0.005	0.095 $\pm$ 0.017	202
Soleus	0.006 $\pm$ 0.002	0.012 $\pm$ 0.001	200

* Significant difference between +/+ and -/- with $P < 0.001$.

Table 3 Carcass analysis of homozygous, heterozygous and wild-type animals

Genotype	Weight (g)		
	Body	Carcass	Fat
+/+ (n = 4)	31.92 $\pm$ 3.63	22.84 $\pm$ 2.41	2.01 $\pm$ 0.46
+/- (n = 5)	32.76 $\pm$ 1.36	23.87 $\pm$ 1.12	2.17 $\pm$ 0.78
-/- (n = 5)	42.28 $\pm$ 2.67*	33.94 $\pm$ 2.07†	2.01 $\pm$ 0.30

Significant difference between +/+ and -/- with $P < 0.01$ (*) and $P < 0.001$ (†).

a 1 MMQKLQMYVY IYLFMLIAAG PVDLNEGSEER EENVEKEGLC NACAWRQNT 50
 51 YSRIEAIKIQ ILSKLRLETA PNISKDAIRQ LLPRAPPLRE LIDQYDVQRD 100
 101 DSSDGSLEDD DYHATTETII TMPTESDFLM QADGKPKCCF KFKSSKIQYN 150
 151 KVVKAQLWIY LRPVKTPPTTV FVQILRLIKP MKDGTRYTGI RSLKLDMSPG 200
 201 TGIWQSIDVK TVLQNLWKQP ESNLGIEIKA LDENGHDLAV TFFPGGEDGL 250
 251 NPFLEVVKVTD TPKR^{*}SR^{*}DDFG LDCDEHSTES RCCRYPLTVD FEAFGWDWII 300
 301 APKRYKANYC^{*} SGECE^{*}FVFLQ KYPTH^{*}LVHQ ANPRGSAGPC^{*} CTPTKMSPIN 350
 351 MLYFNGKEQI IYGKIPAMVV DRCGCS^{*} 376

b

	GDF-8	MIS	inhibin α	inhibin βA	inhibin βB	inhibin βC	GDF-10	BMP-3/osteogenin	BMP-2	BMP-4	BMP-5	Vgr-1 (BMP-6)	OP-1 (BMP-7)	OP-2 (BMP-8)	GDF-5 (CDMP-1)	GDF-6 (CDMP-2)	GDF-7	GDF-1	GDF-3/Vgr-2	GDF-9	nodal	lefty	GDNF	TGF- $\beta 1$	TGF- $\beta 2$	TGF- $\beta 3$
GDF-8	100	31	26	38	42	38	36	38	41	38	42	45	42	40	37	38	37	35	41	27	33	22	16	34	37	37
MIS	-	100	18	24	25	29	31	30	27	27	24	24	27	27	27	26	25	34	22	21	22	15	18	28	23	25
inhibin α	-	-	100	26	25	26	29	22	22	24	25	24	26	24	25	26	23	25	26	23	25	20	23	22	24	
inhibin βA	-	-	-	100	63	53	35	36	42	41	43	44	43	42	40	43	41	37	42	30	39	19	18	41	37	36
inhibin βB	-	-	-	-	100	54	34	37	42	42	37	41	42	37	37	38	36	35	41	30	37	25	22	35	34	37
inhibin βC	-	-	-	-	-	100	30	31	37	36	37	35	38	35	36	36	38	40	34	28	36	20	18	33	34	35
GDF-10	-	-	-	-	-	-	100	83	46	45	42	44	42	40	43	42	42	40	39	28	40	24	19	29	30	29
BMP-3/osteogenin	-	-	-	-	-	-	-	100	48	47	43	44	42	41	47	44	46	42	42	29	41	25	17	32	32	32
BMP-2	-	-	-	-	-	-	-	-	100	92	61	61	60	55	57	54	57	42	53	34	42	21	23	35	34	36
BMP-4	-	-	-	-	-	-	-	-	-	100	59	60	58	55	57	55	57	43	50	35	40	21	22	34	33	35
BMP-5	-	-	-	-	-	-	-	-	-	-	100	91	88	74	52	54	52	46	50	31	40	20	23	34	35	36
Vgr-1 (BMP-6)	-	-	-	-	-	-	-	-	-	-	-	100	87	75	51	53	52	46	53	31	43	21	21	35	37	39
OP-1 (BMP-7)	-	-	-	-	-	-	-	-	-	-	-	-	100	74	51	53	53	47	50	30	41	18	22	34	38	38
OP-2 (BMP-8)	-	-	-	-	-	-	-	-	-	-	-	-	-	100	50	51	53	47	55	27	44	21	24	30	35	38
GDF-5 (CDMP-1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	82	80	46	48	32	42	15	22	33	34	37
GDF-6 (CDMP-2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	82	43	46	32	43	17	22	34	35	37
GDF-7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	48	45	33	45	18	24	36	35	38
GDF-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	50	28	35	18	18	33	32	33
GDF-3/Vgr-2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	32	40	19	24	36	31	32
GDF-9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	31	19	20	22	25	24
nodal	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	18	19	28	32	31
lefty	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	12	25	23	27
GDNF	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	19	17	21
TGF- $\beta 1$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	74	78
TGF- $\beta 2$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	82
TGF- $\beta 3$	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100

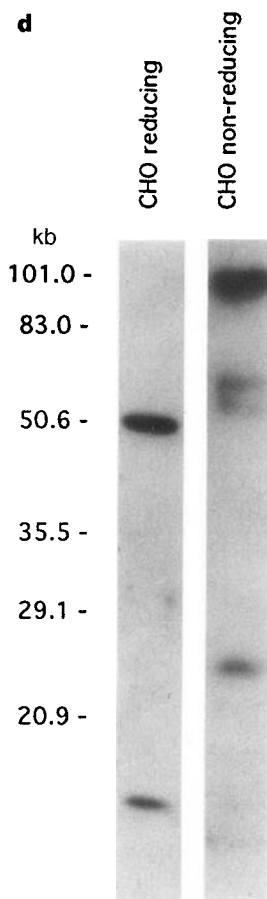
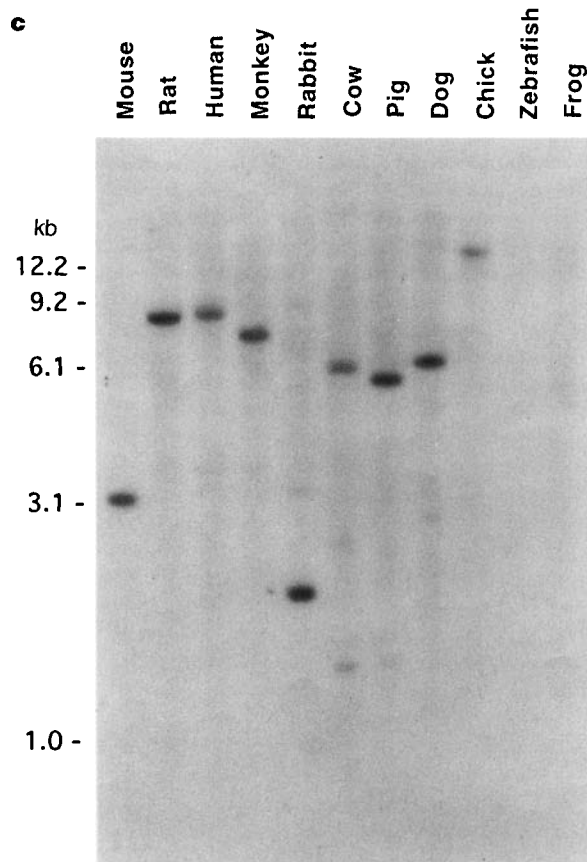
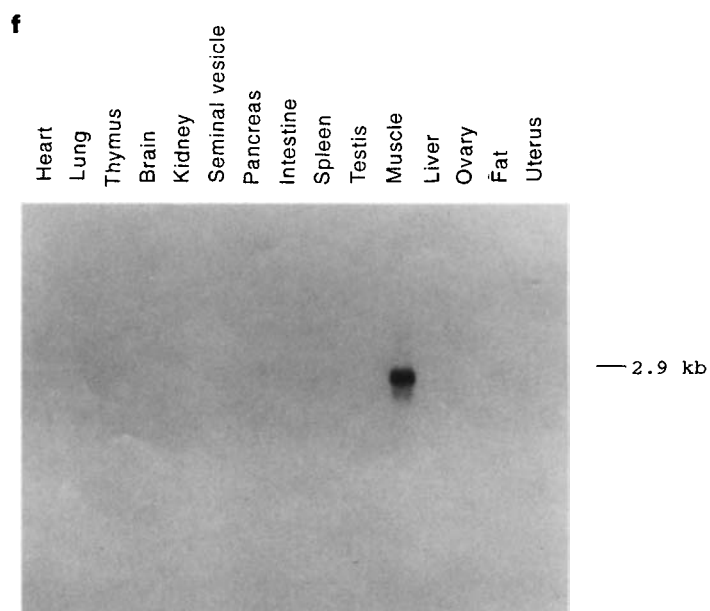
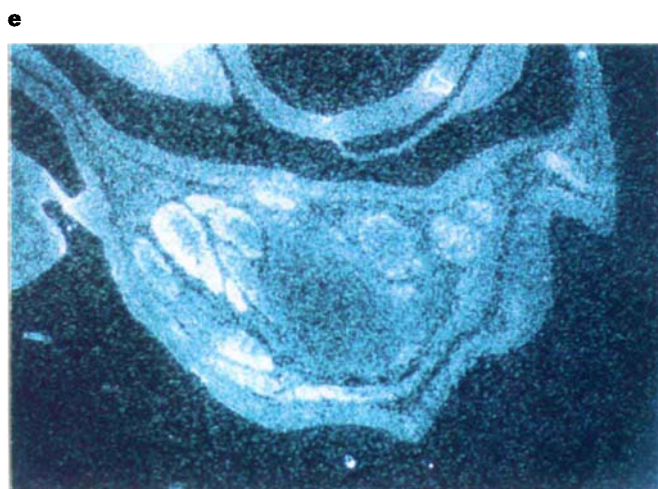
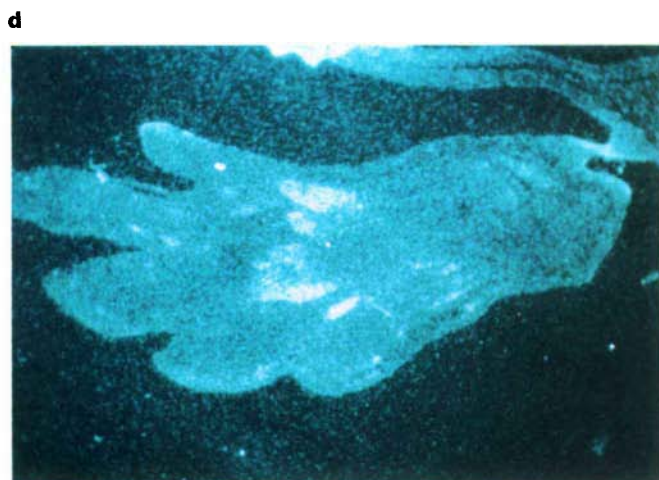
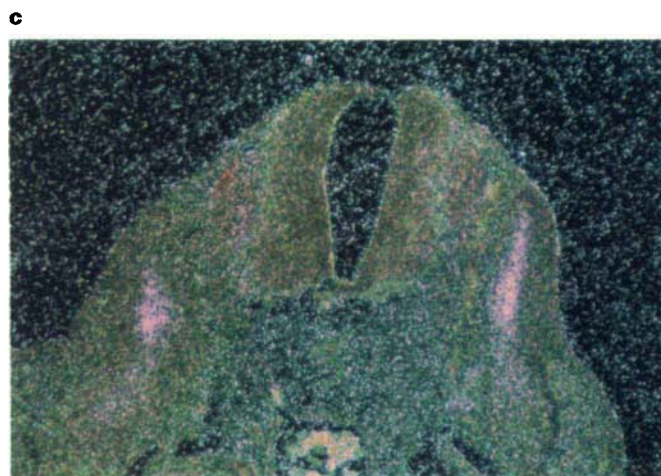


Figure 1 a, Amino-acid sequence of murine GDF-8. The predicted proteolytic processing site is underlined, and the conserved cysteine residues in the C-terminal region are marked by asterisks. The GenBank accession number for GDF-8 is U84005. **b**, Pairwise comparisons between mammalian TGF- β family members. Numbers represent per cent amino-acid identity beginning with the first conserved cysteine residue and extending to the C terminus. Subgroups of highly related sequences are boxed. Human sequences were used for all calculations except for nodal, GDF-3, GDF-7 and lefty, for which murine sequences were used. **c**, Genomic Southern analysis of DNA isolated from different species. **d**, Secretion, processing and dimerization of GDF-8.



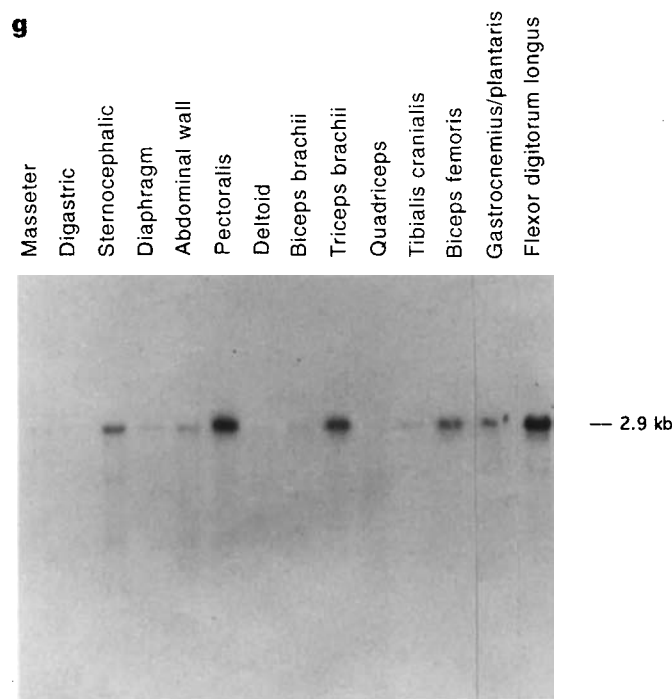


Figure 2 (Opposite and above) Expression of GDF-8 in murine embryonic and adult tissues. **a, b** Whole-mount *in situ* hybridization analysis of day 9.5 p.c. (**a**) and day 10.5 p.c. (**b**) mouse embryos. Hybridization is restricted to developing somites. **c-e**, *In situ* hybridization analysis of sections prepared from paraffin-embedded embryos. **c**, Specific expression of GDF-8 in developing myotome in a horizontal section at the level of the forelimb of a day 10.5 p.c. embryo. **d, e**, Specific expression of GDF-8 in muscles of the developing foot and hip region, respectively, of day 14.5 p.c. embryos. No hybridization was seen using sense probes either on sections or for whole-mount analysis (data not shown). **f, g**, Northern analysis of adult tissues (5 µg of twice-poly(A)⁺-selected RNA) and individual adult muscles (20 µg of total RNA).

($P < 0.001$).

Homozygous mutants were viable and fertile when crossed to C57BL/6 mice and to each other. Homozygous mutant animals, however, were about 30% larger than their heterozygous and wild-type littermates (Table 1). The difference between mutant and wild-type body weights appeared to be relatively constant irrespective of age and sex in adult animals. Adult mutants also displayed an abnormal body shape, with pronounced shoulders and hips. When the skin was removed from animals that had been killed, it was apparent that the muscles of the mutants were much larger than those of wild-type animals (Fig. 4a-d). The increase in skeletal muscle mass appeared to be widespread throughout the body. Individual muscles isolated from homozygous mutant animals weighed about 2-3 times more than those isolated from wild-type littermates (Table 2). Although the magnitude of the weight increase appeared to correlate roughly with the level of GDF-8 expression in the muscles examined (compare Table 2 with Fig. 2g), a more extensive quantification will be necessary to assess the significance of this trend.

To determine whether the increased muscle mass could account for the entire difference in total body weights between wild-type and mutant animals or whether many tissues were generally larger in the mutants, we compared the total body weights to carcass weights. As shown in Table 3, the difference in carcass weights between wild-

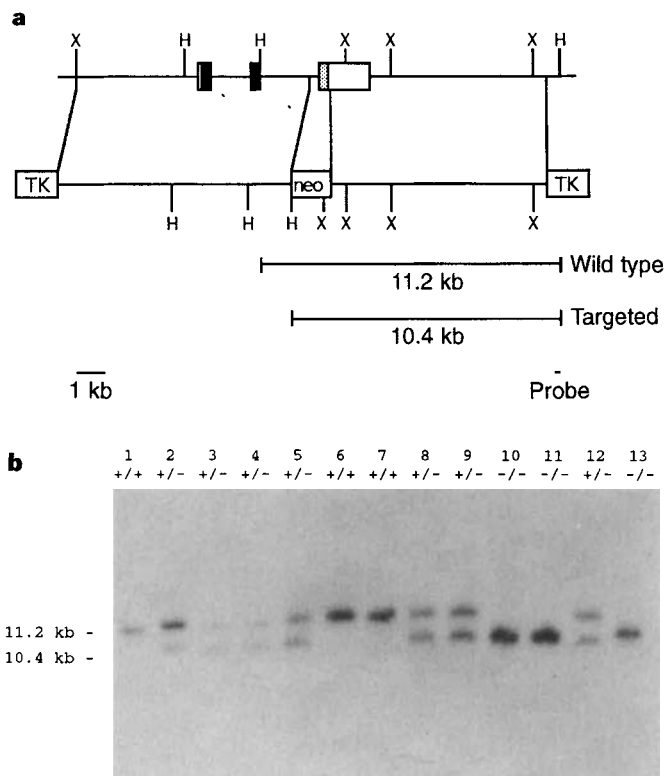


Figure 3 Construction of GDF-8 null mice by homologous targeting. **a**, Map of the GDF-8 locus (top line) and targeting construct (second line). The black and stippled boxes represent coding sequences for the pro- and C-terminal regions, respectively. The white boxes represent 5' and 3' untranslated sequences. The targeting construct contains a total of 17 kb of homology with the GDF-8 gene. A probe derived from the region downstream of the 3' homology fragment and upstream of the most distal *Hind*III site shown hybridizes to an 11.2 kb *Hind*III fragment in the GDF-8 gene and a 10.4 kb fragment in an homologously targeted gene. Abbreviations: H, *Hind*III; X, *Xba*I. **b**, Genomic Southern analysis of offspring derived from a mating of heterozygous mutant mice. DNA prepared from wild-type 129 SV/J mice (lane 1), targeted embryonic stem cells (lane 2), F1 heterozygous mice (lanes 3 and 4), and offspring derived from a mating of these mice (lanes 5-13).

type and mutant animals was comparable to the difference in total body weights. Moreover, because the fat content of mutant and wild-type animals was similar, these data are consistent with all of the total body weight difference resulting from an increase in skeletal muscle mass, although we have not formally ruled out the possibility that differences in bone mass might also contribute to the differences in total body mass.

To determine whether the increase in skeletal muscle mass resulted from hyperplasia or from hypertrophy, we carried out histological analysis of several different muscle groups. As shown in Fig. 4e, the mutant muscle appeared grossly normal. No excess connective tissue or fat was seen and there were no obvious signs of degeneration, such as widely varying fibre sizes (see below) or centrally placed nuclei. Quantification of the number of muscle fibres showed that at the widest portion of the tibialis cranialis muscle, the total cell number was 86% higher in mutant animals compared to wild-type littermates (mutant = $5,470 \pm 121$ ($n = 3$); wild-type = $2,936 \pm 288$ ($n = 3$); $P < 0.01$). Consistent with this result was the finding that the amount of DNA extracted from mutant muscle was roughly 50% higher than from wild-type muscle (mutant = 350 ± 69 µg ($n = 4$); wild-type = 233 ± 54 µg ($n = 3$)) from pooled gastrocnemius, plantaris, triceps brachii, tibialis cranialis and pectoralis muscles; $P = 0.05$). Hence a large part of the increase in skeletal muscle mass resulted from muscle cell hyper-

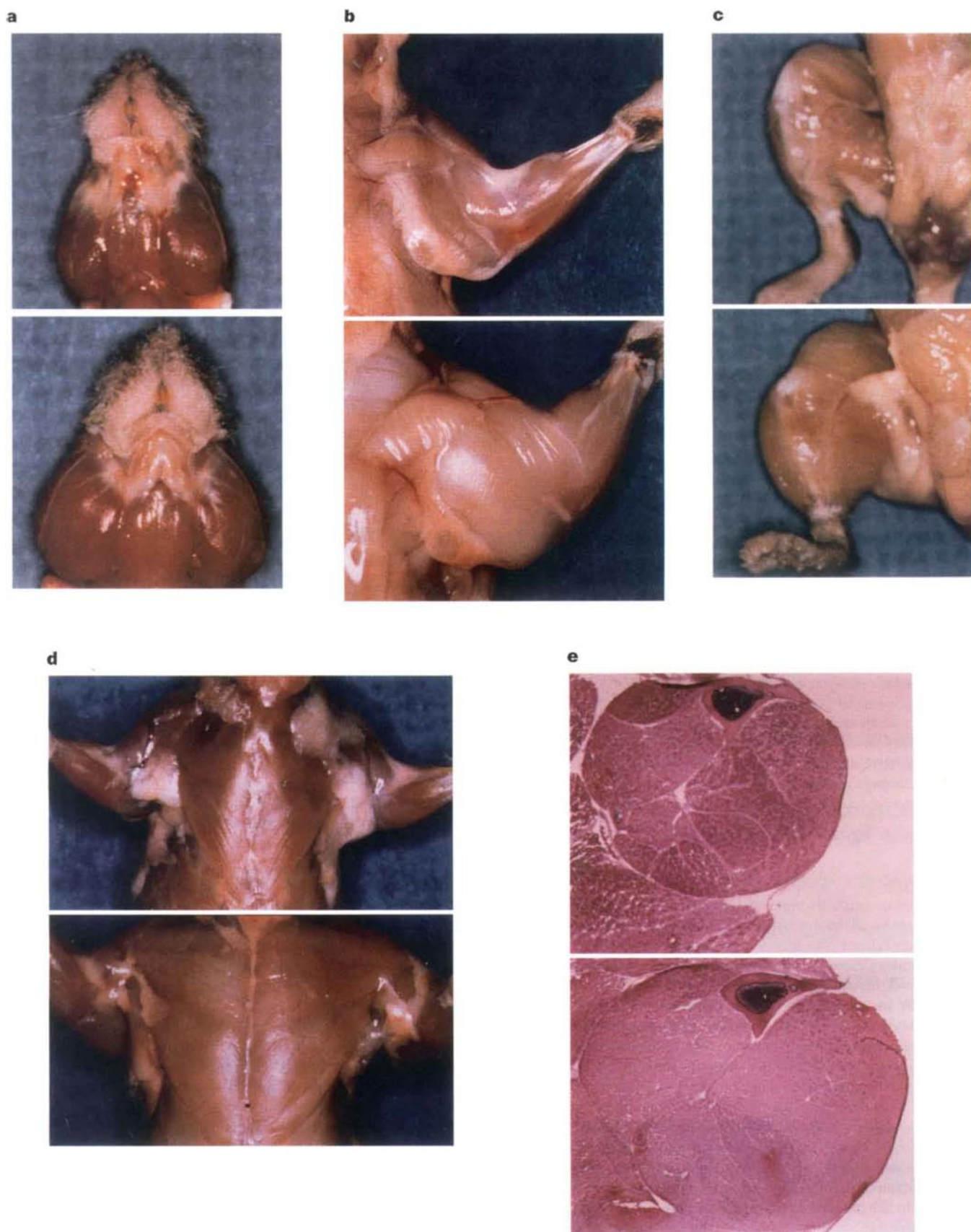


Figure 4 Increased skeletal muscle mass in GDF-8 null mice (bottom panels) compared to wild-type littermates (top panels). **a-d**, Facial (**a**), upper limb (**b**), lower limb (**c**) and pectoral (**d**) muscles of skinned animals. **e**, Sections of distal hindlimbs stained with haematoxylin and eosin.

plasia. However, muscle fibre hypertrophy also appeared to contribute to the overall increase in muscle mass. The mean fibre diameter of the tibialis cranialis muscle and gastrocnemius muscle was 7% and 22% larger, respectively, in mutant animals compared to wild-type littermates, suggesting that the cross-sectional area of the fibres was increased by about 14% and 49%, respectively. Notably, although the mean fibre diameter was larger in the mutants, the standard deviation in fibre sizes was similar between mutant and wild-type muscle (Fig. 5), consistent with the absence of muscle degeneration in mutant animals. The increase in fibre size was also consistent with the finding that the protein to DNA ratio (w/w) was slightly increased in mutant compared to wild-type muscle (mutant = 871 ± 111 ($n = 4$); wild-type = 624 ± 85 ($n = 3$); $P < 0.05$).

In several respects, the biological properties of GDF-8 resemble those of leptin^{2,3}. Both GDF-8 and leptin are secreted molecules that are highly tissue-restricted in their expression patterns, and loss of GDF-8 or leptin function results in increased mass of the tissues that normally produce these factors. Leptin appears to exert its biological effect on adipose tissue indirectly by acting on the central nervous system. Additional experiments will be required to determine whether GDF-8 acts locally or systemically to regulate muscle mass and whether GDF-8 acts directly or indirectly on muscle cells. Elucidating the mechanism of action of GDF-8 will be important for understanding not only the role that GDF-8 plays in regulating muscle development but also the role, if any, that GDF-8 plays in processes such as development of rhabdomyosarcomas, exercise-induced hypertrophy or regeneration following muscle injury. Ultimately, GDF-8 itself or molecules that inhibit GDF-8 signalling may prove useful in the treatment of musculodegenerative states such as muscular dystrophy, neuromuscular diseases, or cancer cachexia. Inhibiting GDF-8 function could also have important applications in agriculture; indeed, as we have shown, GDF-8 homologous sequences appear to be present in various farm animals, including chickens, cows and pigs.

Whatever the precise biological activity of GDF-8 may be, GDF-8 is unique among growth factor-like molecules in that loss of GDF-8 function leads to a dramatic and widespread increase in skeletal muscle mass. On the basis of the phenotype of GDF-8 null mice and the tissue-specificity of GDF-8 expression, we will hereafter refer to GDF-8 as myostatin. □

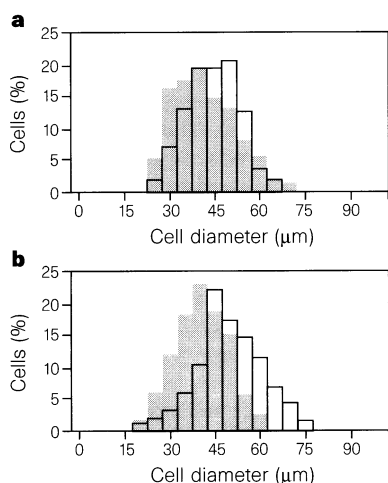


Figure 5 Muscle fibre size distribution in mutant (open bars) and wild-type (shaded bars) animals. Smallest cross-sectional fibre widths were measured for **a**, wild-type ($n = 1,761$) and mutant ($n = 1,052$) tibialis cranialis or **b**, wild-type ($n = 900$) and mutant ($n = 900$) gastrocnemius muscles, and fibre sizes were plotted as a per cent of total fibre number. Standard deviations were 9 and 10 μm , respectively, for wild-type and mutant tibialis cranialis and 11 and 9 μm , respectively, for wild-type and mutant gastrocnemius muscles.

Methods

Cloning of GDF-8. GDF-8 was identified from a mixture of PCR products obtained with primers SJL141 (5'-CCGGAATTCGGITGG(G/C/A)A(G/A/T/C)(A/G)(T/C)TGG(A/G)TI(A/G)TI(T/G)CICC-3'), which encodes the amino-acid sequence GW(H/Q/N/K/D/E)(D/N)W(V/I/M)(V/I/M)(A/S)P, and SJL147 (5'-CCGGAATTC(G/A)CAI(G/C)C(G/A)CA(G/A)CT(G/A/T/C)TCIACI(G/A)(T/C)CAT-3'), the reverse complement of which encodes the amino-acid sequence M(V/I/M/T/A)V(D/E)SC(G/A)C. PCR using these primers was carried out with 2 μg murine genomic DNA at 94 °C for 1 min, 50 °C for 2 min, and 72 °C for 2 min for 40 cycles. PCR products were digested with *Eco*RI, and products about 280 nucleotides long were gel-purified and subcloned into Bluescript (Stratagene, La Jolla, CA). GDF-8 represented 4 out of 110 subclones analysed. A full-length murine GDF-8 cDNA clone was isolated by screening a muscle library using the partial GDF-8 sequence as a probe. The muscle cDNA library was constructed in lambda ZAPII according to the instructions provided by Stratagene and screened without amplification. Library screening and analysis of cDNA clones were carried out as described previously⁴. For genomic Southern analysis of different species, 20 μg DNA were digested with *Eco*RI, electrophoresed, blotted and probed with a C-terminal fragment of GDF-8. All DNA samples were obtained from Clontech (Palo Alto, CA) except for the frog and zebrafish samples, which were kindly provided by D. Brown and M. Halpern, respectively. Filters were hybridized in 5 × SSPE (50 mM sodium phosphate, 5 mM EDTA, 900 mM NaCl), 10% dextran sulphate, 50% formamide, 1% SDS, 200 $\mu\text{g ml}^{-1}$ salmon sperm DNA, and 1% each of ficoll, bovine serum albumin and polyvinylpyrrolidone and washed in 2 × SSC, 1% SDS at 33 °C.

Protein analysis. The GDF-8 coding sequence was cloned into the pMSXND expression vector⁵ and transfected into CHO cells. Following G418 selection, the cells were selected in 0.2 μM methotrexate, and conditioned medium from resistant cells was concentrated and electrophoresed on SDS gels. Conditioned medium was prepared by Cell Trends (Middletown, MD). For preparation of anti-GDF-8 serum, the C-terminal region of GDF-8 (amino acids 268 to 376) was expressed in bacteria using the RSET vector (Invitrogen, San Diego, CA), purified using a nickel chelate column, and injected into rabbits. All immunizations were carried out by Spring Valley Labs (Woodbine, MD). Western analysis using [¹²⁵I]iodoprotein A was carried out as described⁶.

In situ hybridization and northern analysis. For all *in situ* hybridization experiments, probes corresponding to the C-terminal region of GDF-8 were excluded in order to avoid possible crossreactivity with other members of the superfamily. Whole-mount *in situ* hybridization analysis was carried out as described⁷ except that blocking and antibody incubation steps were carried out as in ref. 8. Alkaline phosphatase reactions were carried out for 3 h for day 10.5 embryos and overnight for day 9.5 embryos. Hybridization was carried out using digoxigenin-labelled probes spanning nucleotides 8–811 and 1,298–2,676, which correspond to the pro- region and 3' untranslated regions, respectively. *In situ* hybridization to sections was carried out as described⁹ using ³⁵S-labelled probes ranging from about 100–650 bases in length and spanning nucleotides 8–793 and 1,566–2,595. Following hybridization and washing, slides were dipped in NTB-3 photographic emulsion, exposed for 16–19 days, developed and stained with either haematoxylin and eosin or toluidine blue. RNA isolation, poly(A)⁺ selection, and northern analysis were carried out as described previously⁴.

Construction and analysis of GDF-8 null mice. A murine 129 SV/J genomic library was prepared in lambda FIX II according to the instructions provided by Stratagene. The structure of the GDF-8 gene was deduced from restriction mapping and partial sequencing of phase clones isolated from this library. Vectors for preparing the targeting construct were kindly provided by P. Soriano and K. Thomas. R1 ES cells (kindly provided by A. Nagy, R. Nagy and W. Abramow-Newerly) were transfected with the targeting construct, selected with gancyclovir (2 μM) and G418 (250 $\mu\text{g ml}^{-1}$), and analysed by Southern analysis. Homologously targeted clones were injected into C57BL/6 blastocysts and transferred into pseudopregnant females. Germ-line transmission of the targeted allele was obtained in a total of 9 male chimaeras from 5 independently-derived ES clones. Genomic Southern blots were hybridized at 42 °C as described above and washed in 0.2 × SSC, 0.1% SDS at 42 °C.

For whole-leg analysis, legs of 14-week-old male mice were skinned and treated with 0.2 M EDTA in PBS at 4 °C for 4 weeks followed by 0.5 M sucrose in PBS at 4 °C. For fibre number and size analysis, samples were directly mounted

and frozen in isopentane as described¹⁰. Sections (10–30 µm) were prepared using a cryostat and stained with haematoxylin and eosin. Muscle fibre numbers were determined from sections taken from the widest part of the tibialis cranialis muscle. Muscle fibre sizes were measured from photographs of sections of tibialis cranialis and gastrocnemius muscles.

Carcasses were prepared from shaved 4-month-old mice by removing all of the internal organs and associated fat and connective tissue. Fat content of carcasses was determined as described¹¹. For protein and DNA analysis, muscles from 11–14-week-old males were dissected free of attachments, weighed and homogenized in 150 mM NaCl, 100 mM EDTA. Protein concentrations were determined using the Biorad protein assay. DNA was isolated by adding SDS to 1%, treating with 1 mg ml⁻¹ proteinase K overnight at 55 °C, extracting 3 times with phenol and twice with chloroform, and precipitating with ammonium acetate and EtOH. The samples were then digested with 40 µg ml⁻¹ RNase for 1 h at 37 °C, and following proteinase K digestion and phenol and chloroform extractions, the DNA was precipitated twice with ammonium acetate and EtOH.

Received 13 November 1996; accepted 11 February 1997.

- McPherron, A. C. & Lee, S.-J. in *Growth Factors and Cytokines in Health and Disease* Vol. 1B (eds LeRoith, D. & Bondy, C.) 357–393 (JAI, Greenwich, 1996).
- Friedman, J. M. & Leibel, R. L. Tackling a weighty problem. *Cell* **69**, 217–220 (1992).
- Spiegelman, B. M. & Flier, J. S. Adipogenesis and obesity: rounding out the big picture. *Cell* **87**, 377–389 (1996).
- McPherron, A. C. & Lee, S.-J. GDF-3 and GDF-9: two new members of the transforming growth factor-β superfamily containing a novel pattern of cysteines. *J. Biol. Chem.* **268**, 3444–3449 (1993).
- Lee, S.-J. & Nathans, D. Proliferin secreted by cultured cells binds to mannose-6-phosphate receptors. *J. Biol. Chem.* **263**, 3521–3527 (1988).
- Burnette, W. N. Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**, 195–203 (1981).
- Wilkinson, D. G. in *In situ Hybridization: A Practical Approach* (ed. Wilkinson, D. G.) 75–83 (IRL, Oxford, 1992).
- Knecht, A. K., Good, P. J., Dawid, I. B. & Harland, R. M. Dorsal-ventral patterning and differentiation of noggin-induced neural tissue in the absence of mesoderm. *Development* **121**, 1927–1935 (1995).
- Wilkinson, D. G., Bailes, J. A. & McMahon, A. P. Expression of the proto-oncogene int-1 is restricted to specific neural cells in the developing mouse embryo. *Cell* **50**, 79–88 (1987).
- Brumback, R. A. & Leech, R. W. *Color Atlas of Muscle Histochemistry* (PSG, Littleton, MA, 1984).
- Leshner, A. I., Litwin, V. A. & Squibb, R. L. A simple method for carcass analysis. *Physiol. Behav.* **9**, 281–282 (1972).

Acknowledgements. We thank P. Dunlap and T. Bunton for assistance in maintenance and analysis of mice, Y. Litung for advice on growing ES cells, E. Hsiao for assistance with statistical analysis, D. Nathans for comments on the manuscript and T. Zimmers for suggesting the name myostatin. A.C.M. was supported by training grants from the NIH. This work was supported by research grants from the NIH (to A.M.L.) and the Edward Mallinckrodt Jr Foundation and MetaMorphix Inc. (to S.-J.L.). Under an agreement between MetaMorphix, Inc. and the Johns Hopkins University, A.C.M. and S.-J.L. are entitled to a share of sales royalty received by the University from MetaMorphix. The University, A.C.M. and S.-J.L. also own MetaMorphix stock, which is subject to certain restrictions under University policy. S.-J.L. is a consultant to MetaMorphix. The terms of this arrangement are being managed by the University in accordance with its conflict of interest policies.

Correspondence and requests for materials should be addressed to S.-J.L. (e-mail: sejin_lee@qmail.bs.jhu.edu).

Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese

Sven Enerbäck*†‡, Anders Jacobsson*†‡, Elizabeth M. Simpson*†, Carmen Guerra*, Hitoshi Yamashita*†, Mary-Ellen Harper§ & Leslie P. Kozak*

* The Jackson Laboratory, Bar Harbor, Maine 04609, USA

§ Department of Biochemistry, Faculty of Medicine, University of Ottawa, Ottawa, Ontario K1H 9M5, Canada

† These authors contributed equally to this work.

The mitochondrial uncoupling protein (UCP) in the mitochondrial inner membrane of mammalian brown adipose tissue generates heat by uncoupling oxidative phosphorylation¹. This process protects against cold² and regulates energy balance³.

† Present addresses: University of Göteborg, Dept of Molecular Biology, Medicinareg 9C, S-41390 Göteborg, Sweden (S.E.); The Wenner-Gren Institute, Stockholm University, Dept of Metabolic Research, S-10691 Stockholm, Sweden (A.J.); Dept of Hygiene, National Defense Medical College, 3-2 Namiki, Tokorozawa, Saitama 359, Japan (H.Y.).

Manipulation of thermogenesis could be an effective strategy against obesity^{4–9}. Here we determine the role of UCP in the regulation of body mass by targeted inactivation of the gene encoding it. We find that UCP-deficient mice consume less oxygen after treatment with a β3-adrenergic-receptor agonist and that they are sensitive to cold, indicating that their thermoregulation is defective. However, this deficiency caused neither hyperphagia nor obesity in mice fed on either a standard or a high-fat diet. We propose that the loss of UCP may be compensated by UCP2, a newly discovered homologue of UCP; this gene is ubiquitously expressed and is induced in the brown fat of UCP-deficient mice.

The *Ucp* gene was inactivated by homologous recombination in embryonic stem cells (ES) with a targeting vector that replaced a *Bam*HI/*Bgl*II fragment carrying exon 2 and part of exon 3 with the *neo*^r gene, thereby deleting an essential membrane-spanning domain^{10–12} (Fig. 1). One correctly targeted D3 ES cell clone produced a chimera that transmitted the targeted allele (*Ucp*^{tm1}) to progeny. Homozygous and heterozygous mice were recovered in the proportions expected for a single gene mutation. Northern blot analysis demonstrated that $-/-$ mice lack *Ucp* messenger RNA of the correct size (Fig. 2a). A faint signal from RNA products, both larger and smaller, can be detected on overexposed films; however, the absence of any UCP by immunoblot analysis indicates that this RNA does not contribute to any detectable immunoreactive protein (Fig. 2b). Mice with only one intact copy ($+/-$) of the gene produce *Ucp* mRNA at levels comparable to wild type ($+/+$) mice, as expected for an inducible gene required for thermoregulation. No differences were seen in the levels of transcripts for mitochondrial-localized cytochrome oxidase, the adipocyte-specific aP2 protein (Fig. 2a), cytoplasmic or mitochondrial glycerol-3-phosphate dehydrogenases, the β3 adrenergic receptor or leptin (data not shown).

The brown adipose tissue of UCP-deficient mice has enlarged lipid vacuoles (Fig. 3), but in all other respects it appears normal, including in the shape of its mitochondria (data not shown). The increased lipid deposition is to be expected in the absence of an uncoupling mechanism to dissipate the protonmotive force across the mitochondrial membrane¹³.

Given the importance of UCP in the regulation of energy expenditure in rodents and the ability of β3-adrenergic agonists to stimulate this energy expenditure through their exclusive effects on brown and white adipose tissue¹⁴, we studied oxygen consumption in resting UCP-deficient mice before and after injection of the β3-adrenergic agonist CL 316,243. We found no significant differences in resting oxygen consumption (that is, preinjection values) between wild-type ($+/+$), heterozygous ($+/-$) and UCP-deficient ($-/-$) mice at 28 °C (Fig. 4a). The oxygen consumption after treatment with CL 316,243 was significantly blunted in UCP-deficient mice, however, compared to wild-type and heterozygous controls. Whereas there was an approximate doubling of resting oxygen consumption in the latter two groups, which is normal for rodents¹⁵, in UCP-deficient mice stimulation was only 34%. These results indicate an abnormal response to β3-adrenergic stimulation, like that found for *Ucp*-DTA mice (*Ucp*-diphtheria toxin A chain transgenic mice), when the same protocol and equipment were used⁷.

The effects of UCP deficiency on thermoregulation were assessed by exposing 41 ($-/-$) mice (23 males and 18 females, ranging in age from 27 to 85 days) to cold (5 °C). The time required to lose 10 °C of body heat was used as a measure of sensitivity to cold. Thirty-five ($-/-$) mice (85%) were cold-sensitive and six ($-/-$) mice (15%) were resistant to cold, that is, they maintained body temperature after 24 h in the cold (Fig. 4b). Within the cold-sensitive group, the onset of temperature loss varied from 1.5 to 9.5 hours and was independent of sex or age. The ten heterozygous mice tested were all resistant to cold. These results show that brown-fat thermogenesis is important for protecting both young and

mature mice against cold. Although the role of non-shivering thermogenesis in protecting against cold was expected, the sensitivity of the majority of mature UCP-deficient mice is surprising, contrasting with the resistance to cold of *Ucp*-DTA transgenic mice in which 60–70% of brown adipose tissue had been inactivated by diphtheria toxin^{7,16}.

It is assumed that a role for brown fat in energy balance is determined by UCP-derived thermogenesis. Therefore, as UCP-deficient mice are more sensitive to cold than *Ucp*-DTA mice and consume less oxygen after stimulation with β 3-adrenergic agonists,

the UCP-deficient mice should be as, or more obese than, *Ucp*-DTA mice. We evaluated susceptibility to obesity by feeding mice with either a low-fat laboratory chow or a high-fat diet. UCP-deficient mice showed no significant increase in the development of obesity on the basis of measurements in their body mass, selected fat-pad mass or total body triglyceride content (Table 1). The mass of brown fat was significantly greater in UCP-deficient mice, presumably because of the increased deposition of triglyceride. Food intake was not significantly different in 10 determinations each on groups of 10 male $+/+$, $+/-$ and $-/-$ mice fed a laboratory chow containing 6% fat.

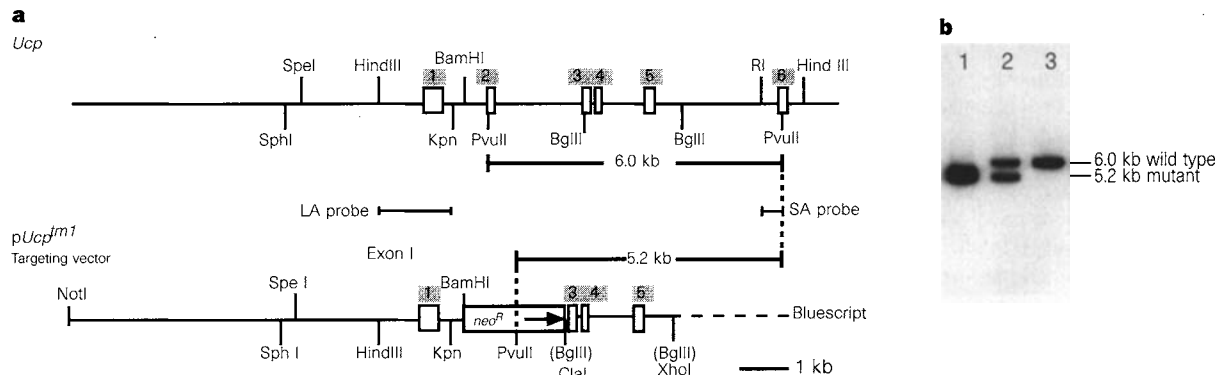


Figure 1 Targeted disruption of the *Ucp* gene. **a**, Restriction enzyme map of the *Ucp* gene showing the relevant sites and the locations of the six exons. **b**, Southern blot analysis of DNA isolated from lane 1, a mouse homozygous for the targeted allele; lane 2, a heterozygous mouse; and lane 3, a mouse homozygous for the wild-type allele. The wild-type and mutated genes gave 6.0- and 5.2-kb

fragments, respectively, after *PvuII* digestion and hybridization with the short-arm (SA) probe. *Bam*HI-cut DNA of the wild-type and targeted genes gave the same predicted-size fragment of 20 kb when hybridized with the long-arm (LA) probe and a probe for the *neor* gene hybridized only to the 5.2-kb *PvuII* fragment (data not shown).

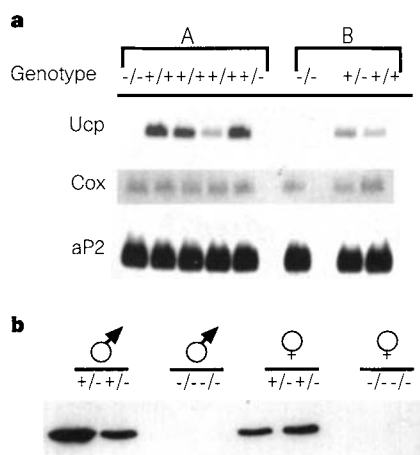


Figure 2 Targeted mice are deficient in *Ucp* mRNA and protein. **a**, Northern blot analysis of brown fat RNA from mice homozygous ($-/-$) and heterozygous ($+/-$) for the targeted allele and wild-type mice ($+/+$). Filters were hybridized sequentially with probes for *Ucp*, cytochrome oxidase (COX) and aP2 mRNA. Animals in A and B were 18 d and 35 d old, respectively. **b**, Immunoblot of UCP, showing the absence of immunoreactive UCP in $-/-$ mice.

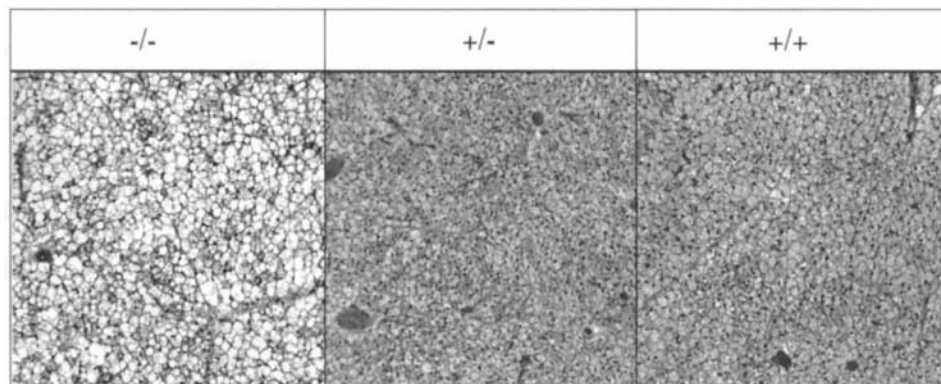


Figure 3 Histology of brown fat shows lipid accumulation in adipocytes. Interscapular brown fat was removed from 3 month-old-male mice of $-/-$,

$+/-$ and $+/+$ genotype at *Ucp* and fixed in Bouin's solution. Paraffin-embedded sections were stained with haematoxylin and eosin.

The absence of obesity in UCP-deficient mice is unexpected. UCP in mammals exists specifically to produce heat: this function provides a mechanism for controlling energy expenditure under conditions of either reduced or enhanced caloric intake and is supported by complex regulation from the central nervous system through neurohormones that modulate feeding behaviour as well as brown-fat thermogenesis¹⁷. The elimination of this system in *Ucp1^{tm1}* mice indicates that an alternative mechanism for maintaining body mass must exist that cannot protect against the cold: the thermogenic glycerol phosphate cycle, which is more prevalent in brown fat than in other tissues¹⁸, or *Ucp2*, a recently described homologue of the brown-fat-specific *Ucp* gene, might be involved.

The UCP2 protein shares 58% overall sequence identity with UCP and their secondary structures, as predicted from hydrophobicity plots, are extraordinarily similar; both have the nucleotide-binding domain. If *Ucp2* can become part of a compensatory mechanism for regulating energy balance in the absence of *Ucp* expression, then its expression might be upregulated in *Ucp1^{tm1}* mice. Using northern blot analysis of RNA, we compared the expression of *Ucp2* in UCP-deficient mice and in normal controls. Whereas *Ucp1* is expressed exclusively in brown fat, *Ucp2* is broadly expressed in different mouse tissues, suggesting that thermogenic mechanisms involving mitochondrial proton leaks may be occurring in many tissues (Fig. 5). *Ucp2* was only upregulated in brown fat (Fig. 5a); this fivefold increase in *Ucp2* mRNA expression in the brown fat of UCP-deficient mice makes it comparable to that in white fat. As

UCP-deficient mice accumulate fat in their brown adipocytes, we investigated whether the induction of *Ucp2* mRNA might be related to this phenotype. We measured *Ucp2* mRNA in the brown adipocytes of transgenic mice overexpressing the gene encoding cytoplasmic glycerol-3-phosphate dehydrogenase, because these mice also accumulate fat in their brown adipocytes⁶. *Ucp2* mRNA was also increased in these mice (data not shown), suggesting that the accumulation of fat in the adipocyte increases *Ucp2* expression. *Ucp2* is the only gene in UCP-deficient mice to show significant alterations in expression; however, this induction may be secondary to the accumulation of lipid in the cells.

The resistance in *Ucp1^{tm1}* mice to developing obesity is striking. This observation is underscored by other mutations affecting brown fat thermogenesis: like *Ucp1^{tm1}* animals, *Ucp*-DTA transgenic mice and mice lacking the β 3-adrenergic-receptor gene both have limited energy expenditure^{7,14}, although it is only the *Ucp*-DTA transgene that abolishes brown fat deposition and causes obesity. A recent study showing that hyperphagia and obesity are absent when *Ucp*-DTA mice are reared at thermoneutrality¹⁹ indicates that even in this model there is no obesity without hyperphagia. Accordingly, the mechanisms that conserve a normal body mass are important in mice, as they are in humans²⁰, so the phenotypes of thermogenesis mutants reflect the efficient maintenance of a normal body mass. As adult humans have very low brown-fat thermogenesis²¹, the UCP-deficient mouse is likely to be a useful model for investigating the relation between energy balance and obesity.

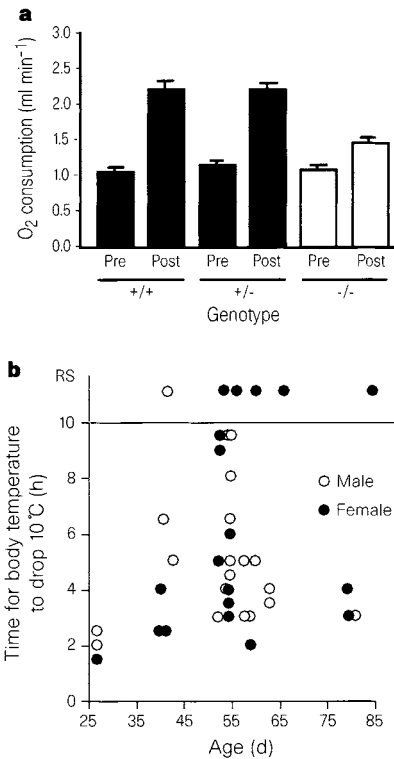


Figure 4 **a**, UCP-deficient mice have normal resting oxygen consumption, but show a blunted response to a β 3-adrenergic agonist. Results are presented as the mean $\pm$ s.e. for the following numbers of male mice: 8 (+/+), 7 (+/-) and 7 (-/-). **b**, Thermoregulation of UCP-deficient mice is defective. Body temperature was measured every 30 min until it was reduced to between 26 to 28°C. Males and females and both immature and mature mice were sensitive to the cold. RS, mice resistant to cold.

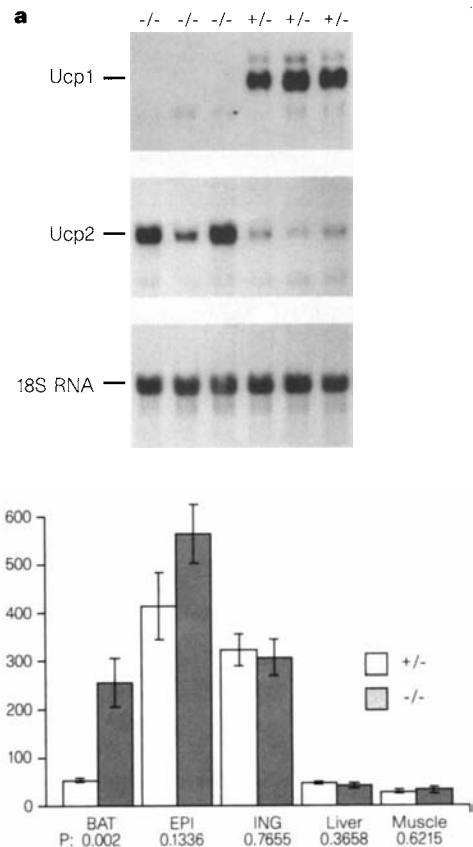


Figure 5 **a**, Northern blot showing that levels of *Ucp2* mRNA are increased in brown adipose tissue of mice homozygous for the allele *Ucp1^{tm1}*. **b**, Bar graph showing that *Ucp2* mRNA is raised only in brown adipose tissue. Data are presented as the mean $\pm$ s.e. of estimates for 6 male mice of 2-3 months old for brown adipose tissue (BAT), epididymal fat (EPI) and inguinal fat (ING), and for 3 mice for liver and muscle. *P* values were determined using the Student's *t*-test with Statview 4.1.

Table 1 Body and fat-pad mass determination indicate that mice homozygous for *Ucp^{tm1}* are not obese

(a) Standard laboratory diet								
Genotype	Sex	Body mass	BAT	RPF	GF	IF		
+/-	M	29.12 ± 2.18	0.10 ± 0.01	0.14 ± 0.02	0.57 ± 0.10	0.37 ± 0.06		
-/-	M	27.91 ± 1.47	0.14 ± 0.02	0.11 ± 0.03	0.40 ± 0.08	0.34 ± 0.06		
		<i>P</i> : 0.6540	0.2035	0.4530	0.1966	0.7481		
+/-	F	21.90 ± 0.86	0.07 ± 0.01	0.07 ± 0.01	0.37 ± 0.07	0.27 ± 0.04		
-/-	F	21.63 ± 0.93	0.11 ± 0.01	0.08 ± 0.01	0.48 ± 0.05	0.28 ± 0.04		
		<i>P</i> : 0.8374	0.0292	0.4342	0.2137	0.8943		
(b) High-fat diet								
Genotype	Sex	Body mass	TG (% body mass)	BAT	RPF	GF	IF	MF
+/-	M	45.77 ± 2.77	34.2 ± 1.7	0.25 ± 0.03	0.46 ± 0.04	2.51 ± 0.24	1.74 ± 0.24	0.98 ± 0.10
-/-	M	37.21 ± 2.99	29.9 ± 3.8	0.28 ± 0.06	0.32 ± 0.04	1.45 ± 0.32	1.09 ± 0.20	0.55 ± 0.10
		<i>P</i> : 0.0918	0.2275	0.5885	0.0649	0.0260	0.1265	0.0259
+/-	F	30.98 ± 1.93	28.4 ± 2.5	0.11 ± 0.01	0.25 ± 0.04	1.24 ± 0.22	0.90 ± 0.14	0.41 ± 0.07
-/-	F	28.92 ± 3.00	23.4 ± 4.6	0.21 ± 0.04	0.22 ± 0.07	1.20 ± 0.45	0.82 ± 0.22	0.38 ± 0.11
		<i>P</i> : 0.5532	0.3151	0.0279	0.7748	0.7207	0.7479	0.8184

Data are expressed as means (in g) ± s.e., except triglyceride (TG), which was assayed with a Sigma kit and is expressed as a percentage of total body mass. *P* values were calculated by using a *t*-test (Statview 4.1). Experiment A contained six mice in each group; Experiment B contained 12 (+/-) males, 5 (-/-) males, 10 (+/-) females and 7 (-/-) females. In **a**, mice were between 68 and 77 d old at the time of killing; in **b**, mice were fed standard laboratory chow until they were 2 months old and then fed on a high-fat diet (5 kcal % in fat; Research Diets, NJ) for three months until they were killed. BAT, brown fat; RPF, retroperitoneal fat; GF, gonadal fat; IF, inguinal fat; MF, mesenteric fat.

Although reduced thermogenesis does not cause obesity, more relevant to the obesity problem is that genetic alterations that enhance the levels of UCP^{8,9} or futile cycling⁶ prevent obesity when mice are either fed a high fat diet or carry an 'obesity' gene. Thus, strategies to reduce obesity should be aimed at enhancing thermogenesis to assist homeostasis during an obesity threat.

Note added in proof. Mice deficient in noradrenaline due to inactivation of dopamine β-hydroxylase fail to induce *Ucp* expression and are cold-sensitive but not obese³¹.

Methods

Gene targeting. The targeting vector, *pUcp^{tm1}*, a derivative of pPGKneobA²², contains an 8-kb fragment of 129/SvJ DNA from the *NotI* site in the multiple cloning site at 5' end of the *Ucp* gene to the *Bam*HI site in intron 1, a *neo^r* gene under the control of the phosphoglycerate kinase gene promoter, a bovine poly(A) signal and a 2-kb *Bgl*II fragment containing exons 3–5 from the *Ucp* gene. Linearized vector was electroporated into D3 embryonic stem cells¹¹ and G418 resistant clones analysed by Southern blot analysis. Correctly targeted clones were microinjected into the blastocysts of C57BL/6J embryos to generate chimaeras that were bred to 129/SvPas mice.

RNA analysis. Total RNA (10 µg), isolated from the brown fat of 35-day-old mice using the guanidinium thiocyanate method²³, was analysed by northern blots²⁴. Hybridization probes were prepared with [³²P]dCTP as described²⁵. Blots were hybridized with probes for the mRNAs of *Ucp*¹², subunit Vb of cytochrome oxidase²⁶, *ap2*²⁷, cytoplasmic and mitochondrial glycerol-3-phosphate dehydrogenase^{18,28} and *Ucp2*. The sequence of an EST (accession number W89820) that encoded a protein with 58% identity to the brown fat specific UCP was submitted to Genbank by the IMAGE Consortium. This cDNA clone, *Ucp2*, was obtained from the American Type Culture Collection and used as a probe to isolate a full length cDNA clones from a brown adipose tissue library¹⁸. Comparisons of the secondary structure of brown fat specific UCP to UCP2 were made with the protein prediction programs of Kyte and Doolittle²⁹ available in GeneWorks.

Immunoblots. Immunoblots of UCP were carried out on mitochondrial fractions separated on SDS–PAGE as described⁸ except that an antirabbit IgG horse radish peroxidase–ECL detection kit (Amersham) was used.

Physiological measurements. Oxygen consumption was measured with computerized equipment including a 1-litre chamber maintained at 28 °C, and air flow of 500 ml per min (regulated with a mass flowmeter; Brooks Instrument Division, Emerson Electric) and an oxygen analyser (Beckman Industrial Oxygen Analyzer, model 755)³⁰. Mice were awake and unrestrained for the study. Resting rate of oxygen consumption was determined when the mouse was curled up and still, about 1.5 h after being placed in the chamber. The β3 agonist, CL 316,243, was dissolved in sterile saline and injected subcutaneously (1 µg per g body weight). Postinjection assessment was made ~1.0 h afterwards when the mouse was still and usually stretched out and panting.

The body temperature of mice was first measured at room temperature with an electronic thermistor equipped with a rectal probe of 1 mm diameter. Mice were then transferred to a cold room maintained at 5 °C where the body temperature was measured every hour until the rectal temperature dropped to 27 °C or for the duration of 24 h if mice were resistant to the cold.

Received 21 November 1996; accepted 6 March 1997.

- Nicholls, D. G. & Locke, R. M. Thermogenic mechanisms in brown fat. *Physiol. Rev.* **64**, 1–64 (1984).
- Foster, D. O. & Frydman, M. L. Tissue distribution of cold-induced thermogenesis in conscious warm- or cold-acclimated rats reevaluated from changes in tissue blood flow: the dominant role of brown adipose tissue in the replacement of shivering by nonshivering thermogenesis. *Can. J. Physiol. Pharmacol.* **57**, 257–270 (1979).
- Rothwell, N. J. & Stock, M. J. A role for brown adipose tissue in diet-induced thermogenesis. *Nature* **281**, 31–35 (1979).
- Champigny, O. *et al.* β3-Adrenergic receptor stimulation restores message and expression of brown-fat mitochondrial uncoupling protein in adult dogs. *Proc. Natl Acad. Sci. USA* **88**, 10774–10777 (1991).
- Himms-Hagen, J. *et al.* Effect of CL-316,243, a thermogenic β3-agonist, on energy balance and brown and white adipose tissues in rat. *Am. J. Physiol.* **266**, R1371–R1382 (1994).
- Kozak, L. P., Kozak, U. C. & Clarke, G. T. Abnormal brown and white fat development in transgenic mice overexpressing glycerol 3-phosphate dehydrogenase. *Genes Dev.* **5**, 2256–2264 (1991).
- Lowell, B. B. *et al.* Development of obesity in transgenic mice after genetic ablation of brown adipose tissue. *Nature* **366**, 740–742 (1993).
- Kopecky, J., Clarke, G., Enerback, S., Spiegelman, B. & Kozak, L. P. Expression of the mitochondrial uncoupling protein gene from the *ap2* promoter prevents genetic obesity. *J. Clin. Invest.* **96**, 2914–2923 (1995).
- Cummings, D. E. *et al.* Genetically lean mice result from targeted disruption of protein kinase A. *Nature* **382**, 622–626 (1996).
- Thomas, K. R. & Capecchi, M. R. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell* **51**, 503–512 (1987).
- Gossler, A., Doetschman, T. C., Korn, R., Serfling, E. & Kemler, R. Transgenesis by means of blastocyst-derived embryonic stem cell lines. *Proc. Natl Acad. Sci. USA* **83**, 9065–9069 (1986).
- Kozak, L. P., Britton, J. H., Kozak, U. C. & Wells, J. M. The mitochondrial uncoupling protein gene. Correlation of exon structure to transmembrane domains. *J. Biol. Chem.* **263**, 12274–12277 (1988).
- Cannon, B. & Vogel, G. The mitochondrial ATPase of brown adipose tissue. Purification and comparison with mitochondria ATPase from beef heart. *FEBS Lett.* **76**, 284–289 (1977).
- Susulic, V. S. *et al.* Targeted disruption of the β3-adrenergic receptor gene. *J. Biol. Chem.* **270**, 29483–29492 (1995).
- Holloway, B. R. *et al.* ICI D7114 a novel selective β-adrenoceptor agonist selectively stimulates brown fat and increases whole body oxygen consumption. *Br. J. Pharmacol.* **104**, 97–104 (1991).
- Hamann, A., Flier, J. S. & Lowell, B. B. Decreased brown fat markedly enhances susceptibility to diet-induced obesity, diabetes and hyperlipidemia. *Endocrinology* **137**, 21–29 (1996).
- Billington, C. J., Briggs, J. E., Grace, M. & Levine, A. S. Effects of intracerebroventricular injection of neuropeptide Y on energy metabolism. *Am. J. Physiol.* **266**, R1765–R1770 (1991).
- Kozak, R. A. *et al.* Sequence and tissue-dependent RNA expression of mouse FAD-linked glycerol-3-phosphate dehydrogenase. *Arch. Biochem. Biophys.* **336**, 97–104 (1996).
- Melnyk, A., Harper, M. E. & Himms-Hagen, J. Raising at the thermoneutrality prevents obesity and hyperphagia in brown adipose tissue-ablated transgenic mice. *Am. J. Physiol.* **272**, R1088–R1093 (1997).
- Leibel, R. L., Rosenbaum, M. & Hirsch, J. Changes in energy expenditure resulting from altered body weight. *N. Engl. J. Med.* **332**, 621–628 (1995).
- Lean, M. E. J. & James, W. P. T. in *Brown Adipose Tissue* (eds Trayhurn, P. & Nicholls, D. G.) 339–365 (Arnold, London, 1986).
- Soriano, P., Montgomery, C., Geske, R. & Bradley, A. Targeted disruption of the *c-src* proto-oncogene leads to osteopetrosis in mice. *Cell* **64**, 693–702 (1991).
- Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analyst. Biochem.* **162**, 156–159 (1987).
- Derman, E. *et al.* Transcriptional control in the production of liver-specific mRNAs. *Cell* **23**, 731–739 (1981).
- Feinberg, A. P. & Vogelstein, B. A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analyst. Biochem.* **132**, 6–13 (1983).
- Zeviani, M. *et al.* Sequence of cDNAs encoding subunit Vb of human and bovine cytochrome c

- oxidase. *Gene* **65**, 1–11 (1988).
27. Bernlohr, D. A., Angus, C. W., Lane, M. D., Bolanowski, M. A. & Kelly, T. J. J. Expression of specific mRNAs during adipose differentiation: identification of an mRNA encoding a homologue of myelin P2 protein. *Proc. Natl Acad. Sci. USA* **81**, 5468–5472 (1984).
 28. Kozak, L. P. & Birkenmeier, E. H. Mouse *sn*-glycerol-3-phosphate dehydrogenase: molecular cloning and genetic mapping of a cDNA sequence. *Proc. Natl Acad. Sci. USA* **80**, 3020–3024 (1983).
 29. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
 30. Ma, S. W. & Foster, D. O. Starvation-induced changes in metabolic rate, blood flow, and regional energy expenditure in rats. *Can. J. Physiol. Pharmacol.* **64**, 1252–1258 (1986).
 31. Thomas, S. A. & Palmiter, R. D. Thermoregulatory and metabolic phenotypes of mice lacking noradrenaline and adrenaline *Nature* **387**, 94–97 (1997).

Acknowledgements. This research was supported by grants from the NIH (L.P.K., E.M.S.), Pfizer (L.P.K.) and NSERC of Canada (M.-E.H.). We thank C. Lerner, G. Sklar and G. Clarke for help in establishing the targeted ES cells, microinjection of embryos, and managing the targeting animals. M.-E.H. thanks J. Himms-Hagen for the use of her indirect calorimetry equipment.

Correspondence and requests for materials should be addressed to L.P.K. (e-mail: lpk@aretha.jax.org).

Thermoregulatory and metabolic phenotypes of mice lacking noradrenaline and adrenaline

Steven A. Thomas & Richard D. Palmiter

Howard Hughes Medical Institute, Department of Biochemistry, Box 357370, University of Washington, Seattle, Washington 98195-7370, USA

Adrenaline and noradrenaline, the main effectors of the sympathetic nervous system and adrenal medulla, respectively, are thought to control adiposity and energy balance through several mechanisms. They promote catabolism of triglycerides and glycogen¹, stimulate food intake when injected into the central nervous system², activate thermogenesis in brown adipose tissue^{3,4}, and regulate heat loss through modulation of peripheral vasoconstriction and piloerection¹. Thermogenesis in brown adipose tissue occurs in response to cold and overeating (diet induced)^{5–7}, and there is an inverse relationship between diet-induced thermogenesis and obesity both in humans⁸ and in animal models^{9–12}. As a potential model for obesity, we generated mice that cannot synthesize noradrenaline or adrenaline by inactivating the gene that encodes dopamine β -hydroxylase. These mice are cold intolerant because they have impaired peripheral vasoconstriction and are unable to induce thermogenesis in brown adipose tissue through uncoupling protein (UCP1). The mutants have increased food intake but do not become obese because their basal metabolic rate is also elevated. The unexpected increase in basal metabolic rate is not due to hyperthyroidism, compensation by the widely expressed uncoupling protein UCP2, or shivering.

Most mice lacking the gene for dopamine β -hydroxylase (*dbh*^{-/-} mice) die *in utero* but can be rescued by supplying the precursor L-threo-3,4-dihydroxyphenylserine (DOPS) in the maternal drinking water from embryonic day 9.5 until birth¹³. Most *dbh*^{-/-} mice become viable adults without further treatment. There is a delay of 1–2 weeks in the rapid adolescent growth phase of the mutant mice¹³, resulting in slightly lower body weight relative to controls (89%, $P = 0.003$; Fig. 1a) at 4–5 months of age, when all studies reported here were performed. The reduction in body weight is distributed equally among the major organs except for the spleen, which was ~10% larger ($P = 0.04$), and brown adipose tissue, which was over twice the weight of controls ($P = 2 \times 10^{-6}$; Fig. 1a). White adipose tissue did not differ between genotypes on a per-animal or per-body weight basis ($P = 0.36$ and $P = 0.85$, respectively; Fig. 1a). In addition, plasma leptin levels of *dbh*^{-/-} mice

(6.43 ± 1.43 ng ml⁻¹, $n = 10$) were similar to those of *dbh*^{+/+} mice (6.64 ± 1.37 ng ml⁻¹, $n = 10$). The increase in brown adipose tissue was not due to an increase in cell number, as measured by DNA content, and the RNA/DNA ratio was normal (Fig. 1a). Histologically, the lipid vacuoles in the brown adipocytes of mutant mice were markedly enlarged compared with those of controls (Fig. 1b), suggesting that the tissue was hypoactive.

The thermogenic activity of brown adipose tissue depends largely on UCP1, which generates heat by dissipating the mitochondrial proton gradient^{3,4}. The expression of UCP1 is regulated transcriptionally by β -adrenergic and thyroid-hormone stimulation¹⁴. There was a large increase in UCP1 mRNA in control mice at room temperature (22°C) relative to thermoneutrality (30°C), and placing them at 4°C for 2 h did not increase UCP1 mRNA further (Fig. 2a), although at longer times it would (data not shown). In contrast, *dbh*^{-/-} mice exhibited low UCP1 mRNA levels under all three conditions (Fig. 2a). Those *dbh*^{-/-} mice with noradrenaline restored by prior administration of DOPS had normal levels of UCP1 mRNA after 2 h at 4°C. Adrenaline is not restored, and the elevated levels of dopamine persist with the regimen of DOPS used here. A single injection of DOPS normalizes UCP1 mRNA levels

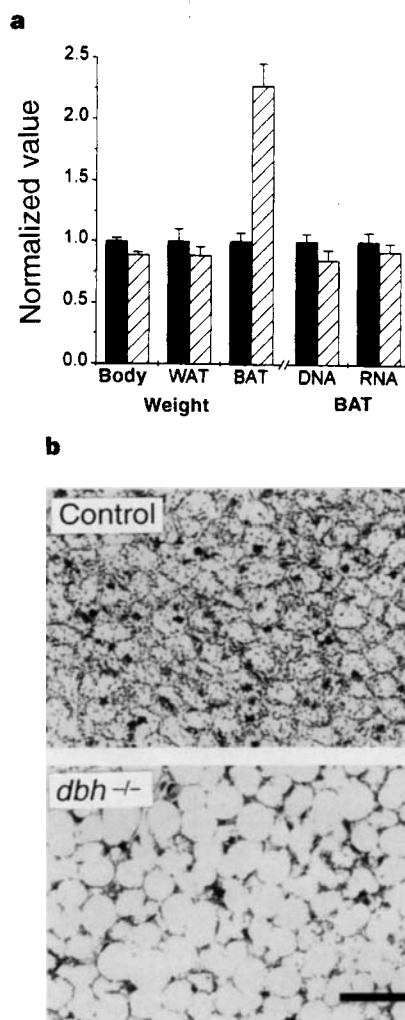


Figure 1 Hypertrophy of brown adipose tissue (BAT) in *dbh*^{-/-} mice (hatched bars) relative to controls (black bars). **a**, BAT weight was selectively increased in the mutant mice despite having normal DNA content (DNA) and RNA/DNA (RNA). Control values normalized to 1 ($n = 11$ for each genotype) were: 31.2 g (body), 599 mg (WAT), 55 mg (BAT), 90 μ g (DNA), 1.66 μ g (RNA/ μ g DNA). **b**, Lipid vacuoles were enlarged in the BAT of the mutant mice. Scale bar, 50 μ m.

within 12 h (data not shown), indicating that the cells in the brown adipose tissue of the mutant mice are indeed brown adipocytes. A role for noradrenaline in the development of brown adipose tissue has been hypothesized because it stimulates proliferation of brown adipocytes¹⁵. Our results suggest, however, that noradrenaline is not essential for development of the normal number of brown adipocytes, but is essential for the normal function of brown adipose tissue.

Because the sympathetic nervous system regulates heat loss as well as thermogenesis, core body temperature was monitored while the mice were housed at 4 °C. The temperature of control mice dropped by about 2 °C in the first 30 min with no further drop at 1 h (Fig. 2b). In contrast, mutant mice steadily lost heat such that at 1 h their core body temperature had dropped by about 13 °C, and most of them became lethargic and shivered. This rate of heat loss is half that of mice killed just before placement at 4 °C (Fig. 2b). By comparison, the body temperature of transgenic mice deficient in brown adipose tissue dropped by only 1.7 °C after 30 h at 4 °C (ref. 10), whereas that of most mice with a targeted disruption of the UCP1 gene dropped by 10 °C in 1.5–9.5 h (ref. 16). These observations suggest that mechanisms other than just brown adipose tissue thermogenesis are impaired in the *dbh*^{-/-} mice. Because adrenaline and noradrenaline regulate vascular resistance¹, a failure of peripheral vasoconstriction could explain the greater heat loss of the mutant mice in the cold. To test this, peripheral blood flow in the tail was recorded before and during the transition into the cold. The mutant mice exhibited blunted vasoconstriction, decreasing blood flow to 65 ± 13% (*n* = 7) of that measured at 22 °C (Fig. 2d). In contrast, control mice decreased blood flow to 15 ± 3% (*n* = 7) of that measured at 22 °C (Fig. 2c); the difference between genotypes was significant (*P* = 0.002). Furthermore, piloerection was reduced in the mutant mice relative to controls.

Consistent with having hypoactive thermogenesis in brown adipose tissue, *dbh*^{-/-} mice have a lower core body temperature when housed at 22 °C (*P* = 10⁻¹¹, Fig. 3a). We therefore anticipated

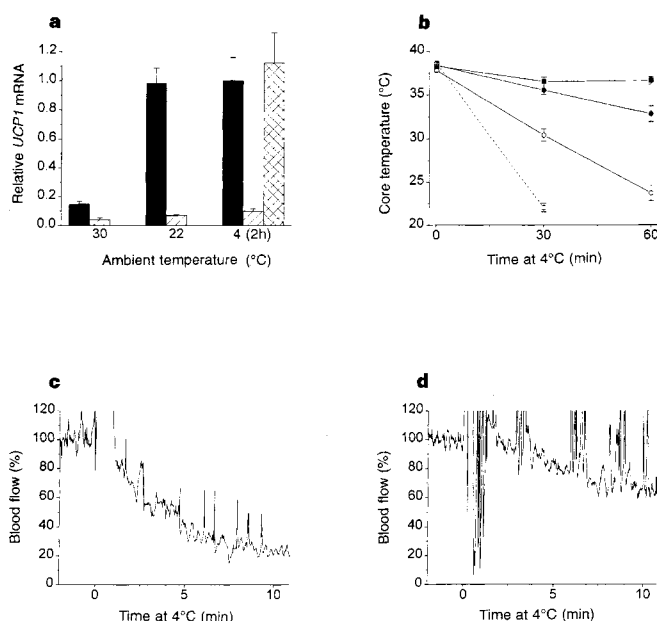


Figure 2 Thermoregulation during cold stress. **a**, UCP1 mRNA levels in brown adipose tissue (BAT) after two weeks at 30 °C, chronically at 22 °C, after 2 h at 4 °C, and after 3 injections of DOPS in *dbh*^{-/-} mice (cross-hatched bars, *n* = 7–11 for each point). UCP1 mRNA levels were normalized to 1.0 for *dbh*^{+/+} mice (black bars) kept at 4 °C for 2 h; **b**, core body temperature of control mice (*n* = 11), *dbh*^{-/-} mice (white open circles, *n* = 17), *dbh*^{-/-} mice injected with DOPS for 5 days (filled circles, *n* = 10), and mice killed just before placement at 4 °C (open squares, *n* = 3); control, filled squares. **c**, **d**, Peripheral blood flow in the tail of a *dbh*^{-/-} mouse (**c**) and *dbh*^{-/-} mouse (**d**) before and after placement at 4 °C from 22 °C. Time 0 denotes when the transition into the cold was made. The mean blood flow at 22 °C was normalized to 100%. Spikes in blood flow are movement artefacts.

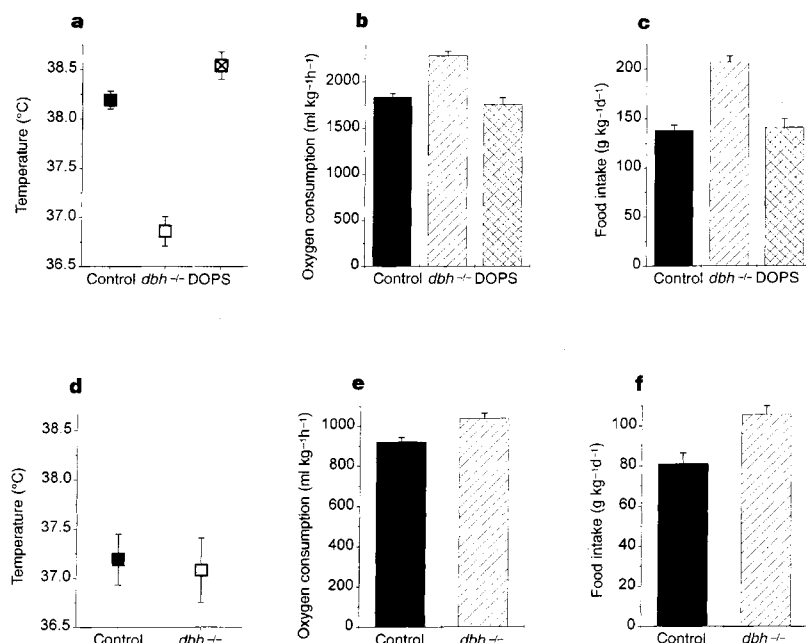


Figure 3 Metabolism and food intake at 22 °C (**a–c**) and thermoneutrality (30 °C, **d–f**). **a**, **d**, Core body temperature. **b**, **e**, Basal oxygen consumption. **c**, **f**, Food intake. Mice (8–10) were used for each measurement except temperature at 22 °C, where 46 mice of each genotype were used for control and *dbh*^{-/-}. Oxygen

consumption was monitored beginning 30 min after a single injection of DOPS; food and water intake were then measured over the next 5 days with DOPS being injected every 12 h. At the end of the 5 days, core body temperature was recorded.

that the basal metabolic rate would be lower in the mutant mice; however, it was elevated by 24% ($P = 10^{-7}$, Fig. 3b). Food intake was also significantly increased in $dbh^{-/-}$ mice ($P = 10^{-7}$, Fig. 3c), possibly in response to the higher metabolic demand. To test whether these differences are due to the physiological absence of noradrenaline, the parameters were measured in the mutant mice following repletion of noradrenaline by DOPS. Importantly, basal metabolic rate of $dbh^{-/-}$ mice was normal several hours after a single injection of DOPS, and body temperature and food intake also normalized in mutants repeatedly injected with DOPS (Fig. 3). Because this increased metabolic rate might be a compensatory response for greater heat loss and lack of brown adipose tissue thermogenesis in the mutants, the above parameters were also measured in mice housed for 2 weeks at 30 °C. In that environment, there was no difference in core body temperature between the genotypes, but basal metabolic rate and food intake remained significantly elevated in the mutant mice relative to controls ($P < 0.004$ for each, Fig. 3). These observations indicate that increased metabolic rate is not due solely to the need to generate additional heat in the mutant mice. Note also that the basal metabolic rate and food consumption for both genotypes maintained at 30 °C were about half the values at 22 °C (Fig. 3).

We tested three potential explanations for the increased metabolic rate of the mutant mice. First, electromyograms were recorded from mutant and control mice to determine whether the mutants chronically shiver in an effort to maintain core body temperature. There was no difference between the genotypes with regard to the standard deviation ($dbh^{+/+}$, $34 \pm 2.8 \mu V$; $dbh^{-/-}$, $34.5 \pm 4.6 \mu V$; $n = 9$ for each) or peak to peak signal ($dbh^{+/+}$, $199 \pm 19 \mu V$; $dbh^{-/-}$, $195 \pm 28 \mu V$) of the EMG. Second, because basal metabolic rate is regulated by the thyroid hormone axis¹⁷, serum levels of thyroxine (T_4), triiodothyronine (T_3) and thyroid-stimulating hormone (TSH) were quantified by radioimmunoassay. There was no significant difference between genotypes with regard to total T_3 , total T_4 , free T_4 or TSH (Table 1). Further evidence that elevated metabolic rate was not the result of hyperthyroidism was provided by our observation that restoration of noradrenaline in the $dbh^{-/-}$ mice depressed basal metabolic rate within several hours of administration, while effects mediated by thyroid hormone typically have a latency of 6–12 h (ref. 17). Finally, a second uncoupling protein, UCP2, has recently been cloned¹⁸ that shares homology with brown adipose tissue-specific UCP1 but is expressed in most tissues. Because UCP2 might be upregulated under conditions of increased heat loss and/or diminished UCP1 expression, levels of UCP2 mRNA were measured. The differences in UCP2 mRNA levels from various tissues were small or absent between the genotypes (Table 2). These results indicate that shivering, hyperthyroidism and increased expression of UCP2 are not responsible for the elevated basal metabolic rate of $dbh^{-/-}$ mice.

The increased metabolic rate in the $dbh^{-/-}$ mice may explain why these mice do not become obese, even though their brown adipose tissue is metabolically hypoactive and their food intake is increased. Because induction of UCP1 is thought to be the main thermogenic mechanism activated by the sympathetic nervous system, and inactivation of the UCP1 gene does not lead to obesity¹⁶, these observations suggest that neither activity of the sympathetic nervous system nor UCP1 are necessary for the regulation of energy stores. Potential sites at which noradrenaline deficiency might cause

increased metabolic rate are the hypothalamic paraventricular nucleus and the sulcal prefrontal cortex, because injection of noradrenaline or clonidine (an α_2 -adrenergic receptor agonist) into either of these sites decreases basal metabolic rate independent of motor activity, although the mechanism is unclear^{19,20}. The elevation in food intake in $dbh^{-/-}$ mice is probably secondary to the increase in basal metabolic rate rather than vice versa because restoration of noradrenaline returns metabolic rate to normal within several hours in the absence of food. Previous studies have demonstrated that injection of noradrenaline or clonidine into the hypothalamic paraventricular nucleus of rats stimulates food intake² and acute pharmacological depletion of noradrenaline reduces food intake²¹. Furthermore, a hypothesis for the initiation and termination of feeding has been elaborated that depends critically on episodic outflow of the sympathetic nervous system to stimulate brown adipose tissue²². Our results with the $dbh^{-/-}$ mice indicate that noradrenaline is not required in either the central or the sympathetic nervous system for feeding. These results are analogous to those obtained with mice lacking neuropeptide Y²³, a transmitter that also stimulates feeding when injected into the hypothalamic paraventricular nucleus²⁴. If compensatory mechanisms for the role of noradrenaline in feeding arose during the development of $dbh^{-/-}$ mice, then restoration of noradrenaline using DOPS might be expected to augment food intake rather than decrease it.

Because $dbh^{-/-}$ mice have normal adiposity, defects in noradrenergic tone may not be a primary cause of obesity. However, the loss of noradrenergic signals at sites in addition to adipose tissue may prevent obesity in $dbh^{-/-}$ mice. Our findings suggest that noradrenaline is important both for energy conservation in the cold and for the regulation of basal metabolic rate, which is elevated in $dbh^{-/-}$ mice housed at thermoneutrality. □

Methods

Animals and histology. The mice used throughout this study were fifth- to seventh-generation male hybrids of 129/SvCPJ and C57BL/6J strains aged 4–5 months. The $dbh^{-/-}$ mice were used as controls because they have normal levels of noradrenaline and adrenaline (S.A.T., A. M. Matsumoto and R.D.P., unpublished observations). Mice were initially housed at 22 °C, and some were then housed at 30 °C for 2 weeks or 4 °C for 2 hours before being killed. The $dbh^{-/-}$ mice rescued with DOPS received subcutaneous injections of 1 mg per g body weight DOPS in a 20 mg ml⁻¹ 0.2 M HCl solution that was neutralized with NaOH just before injection. Injections were given 12 h apart, the last injection being ~5 h before testing. Histology was performed by cardiac perfusion with neutral-buffered formalin (NBF) followed by post-fixation with NBF, dehydration and embedding in paraffin. Sections were stained with haematoxylin and eosin. For data analysis, all values are expressed as mean $\pm$ s.e., with Student's *t*-test used for statistical comparison.

Molecular biology. For white adipose tissue the reproductive and retroperitoneal fat pads were measured. For brown adipose tissue the two interscapular brown fat pads were used. Total nucleic acid (TNA) was isolated by SDS/proteinase K digestion followed by phenol/chloroform extraction and ethanol precipitation. TNA was dissolved and total content determined by absorbance at 260 nm; DNA was determined by fluorescence of H33258; RNA was calculated from TNA minus DNA. To quantify relative levels of UCP1 mRNA, a ³²P-labelled antisense oligonucleotide (5'-GCAGAGCCACCTGGG CTAGGTAGTGC) to the 5'-untranslated region of UCP1 mRNA was hybridized with 10–40 μ g of brown adipose tissue nucleic acid for 12 h at 45 °C, then

Table 1 Serum thyroid hormone and TSH levels

	Total T_3 (ng ml ⁻¹)	Total T_4 (ng ml ⁻¹)	Free T_4 (pg ml ⁻¹)	TSH (ng ml ⁻¹)
$dbh^{+/+}$	1.80 ± 0.15 (12)	28.8 ± 1.9 (12)	37.9 ± 4.1 (10)	484 ± 142 (9)
$dbh^{-/-}$	1.65 ± 0.11 (10)	29.5 ± 1.7 (12)	48.7 ± 5.4 (9)	360 ± 60 (9)
<i>P</i>	0.47	0.77	0.12	0.43

All values are mean $\pm$ s.e. (number of samples).

Table 2 UCP2 mRNA levels

	Brown adipose tissue	White adipose tissue	Heart	Skeletal muscle	Liver
$dbh^{+/+}$	13.2 ± 1.2 (11)	5.60 ± 0.95 (9)	8.5 ± 0.73 (9)	2.45 ± 0.47 (9)	1.78 ± 0.25 (9)
$dbh^{-/-}$	16.8 ± 1.1 (11)	5.75 ± 0.85 (9)	6.6 ± 0.56 (9)	2.45 ± 0.18 (9)	1.82 ± 0.36 (9)
<i>P</i>	0.03	0.90	0.04	1.00	0.92

All values are mean $\pm$ s.e. in c.p.m. μ g⁻¹ TNA (number of samples).

digested with S_1 nuclease, precipitated with trichloroacetic acid, and passed over GF/C filters (Whatman) for scintillation counting, normalized to amount of TNA. For UCP2 mRNA detection, a 360-base pair *Bgl*II-*Sma*I fragment of mouse UCP2 cDNA (Genbank accession no. U69135) was subcloned into Bluescript; an antisense riboprobe was synthesized using T3 RNA polymerase and [α - 32 P]-UTP. The hybridization and detection protocol was similar to that used for UCP1, except hybrids were digested using RNases A and T1 before acid precipitation.

Hormones. Plasma leptin levels were determined by radioimmunoassay (Linco, St Charles, MO)²⁵. Serum thyroid hormone levels were determined by radioimmunoassay (Phoenix, Everett, WA) and mouse TSH levels were determined by radioimmunoassay (AniLytics, Gaithersburg, MD) using the mouse TSH preparation of the National Hormone and Pituitary Institute as a standard.

Physiology. Core body temperature was monitored using a model 43 TA thermometer and model 402 colonic probe (YSI, Yellow Springs, OH). Oxygen consumption was measured using an Oxymax indirect calorimeter (Columbus, OH). Airflow to the chamber was 0.75 l min^{-1} ; samples were taken every minute, with room air reference taken every 30 min. Mice were kept in the chamber for 4–5 h without food or water during the middle of the light cycle. Basal oxygen consumption was determined for each mouse by averaging minimum plateau regions (typically 50–100 data points total), which corresponded to periods of inactivity or sleep. For mice housed at 30°C the Oxymax chamber was also maintained at 30°C . Food and water values are the average daily intake over a 5-day period. For DOPS rescue, oxygen consumption was measured beginning ~ 1 h after a single injection; food and water intake were measured over the next 5 days with injections of DOPS every 12 h; temperature was measured at the end of the 5 days, just before placement at 4°C .

Peripheral blood flow was assessed using a floLAB Server laser Doppler perfusion monitor and a two-fibre right-angle optical probe (Moor, Millway, UK). Mice were paralysed with an intraperitoneal injection of D-tubocurarine HCl (Sigma, St Louis, MO) at $0.7 \mu\text{g per g}$ body weight. About 10 min after injection, the tail was positioned over the probe and lightly fixed in place with tape. The probe and tail were held in place by a custom wooden adapter $\sim 2 \times 2 \times 1$ cm. The probe was repositioned if blood flow was too low for analysing vasoconstriction. Flow was monitored until a stable baseline was achieved, then the whole apparatus was moved to the cold room (4°C) and left there until a new stable baseline was achieved, usually after about 10 min.

EMGs were recorded by placing mice in a restraining tube and inserting two stainless steel safety pins subcutaneously over the left and right back muscles. The electrical signal was passed through a model M-707 Microprobe amplifier (World Precision Instruments, New Haven, CT) and boosted 121 times through two low-pass filters (Frequency Devices, Model 902). The data was filtered at 1 kHz and collected digitally using an INDEC (Sunnyvale, CA) acquisition board, a computer and Fastlab software.

Received 19 December 1996; accepted 26 March 1997.

1. Landsberg, L. & Young, J. B. Catecholamines and the adrenal medulla. In *Williams Textbook of Endocrinology* 8th edn (eds Wilson, J. D. & Foster, D. W.) 621–705 (W.B. Saunders, Philadelphia, 1992).
2. Goldman, C. K., Marino, L. & Leibowitz, S. F. Postsynaptic α_2 -noradrenergic receptors mediate feeding induced by paraventricular nucleus injection of norepinephrine and clonidine. *Eur. J. Pharmacol.* **115**, 11–19 (1985).
3. Nicholls, D. G. & Locke, R. M. Thermogenic mechanisms in brown fat. *Physiol. Rev.* **64**, 1–64 (1984).
4. Klaus, S., Castella, L., Bouillard, F. & Ricquier, D. The uncoupling protein UCP: a membranous mitochondrial ion carrier exclusively expressed in brown adipose tissue. *Int. J. Biochem.* **23**, 791–801 (1991).
5. Rothwell, N. J. & Stock, M. J. A role for brown adipose tissue in diet-induced thermogenesis. *Nature* **281**, 31–35 (1979).
6. Brooks, S. L. *et al.* Increased proton conductance pathway in brown adipose tissue mitochondria of rats exhibiting diet-induced thermogenesis. *Nature* **286**, 274–276 (1980).
7. Glick, Z., Teague, R. J. & Bray, G. A. Brown adipose tissue: thermogenic response increased by a single low protein, high carbohydrate meal. *Science* **213**, 1125–1127 (1981).
8. Jung, R. T., Shetty, P. S., James, W. P. T., Barrand, M. A. & Callingham, B. A. Reduced thermogenesis in obesity. *Nature* **279**, 322–323 (1979).
9. Bray, G. A., York, D. A. & Flier, J. S. Experimental obesity: a homeostatic failure due to defective nutrient stimulation of the sympathetic nervous system. *Vitam. Horm.* **45**, 1–125 (1989).
10. Lowell, B. B. *et al.* Development of obesity in transgenic mice after genetic ablation of brown adipose tissue. *Nature* **366**, 740–742 (1993).
11. Hamann, A., Flier, J. S. & Lowell, B. B. Decreased brown fat markedly enhances susceptibility to diet-induced obesity, diabetes, and hyperlipidemia. *Endocrinology* **137**, 21–29 (1996).
12. Susulic, V. S. *et al.* Targeted disruption of the β_3 -adrenergic receptor gene. *J. Biol. Chem.* **270**, 29483–29492 (1995).
13. Thomas, S. A., Matsumoto, A. M. & Palmiter, R. D. Noradrenaline is essential for mouse fetal development. *Nature* **374**, 643–646 (1995).
14. Silva, J. E. Full expression of uncoupling protein gene requires the concurrence of norepinephrine and triiodothyronine. *Mol. Endocrinol.* **2**, 706–713 (1988).
15. Goleen, A., Collet, A. J., Guay, G. & Bukowiecki, L. J. β -adrenergic stimulation of brown adipocyte

- proliferation. *Am. J. Physiol.* **254**, C175–C182 (1988).
16. Enerback, S. *et al.* Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese. *Nature* **387**, 90–94 (1997).
17. Larson, P. R. & Ingbar, S. H. The thyroid gland. In *Williams Textbook of Endocrinology* 8th edn 357–488 (eds Wilson, J. D. & Foster, D. W.) (W.B. Saunders, Philadelphia, 1992).
18. Fleury, C. *et al.* Uncoupling protein-2: a novel gene linked to obesity and hyperinsulinemia. *Nature Genet.* **15**, 269–273 (1997).
19. Sivi, S. M., Kritikos, A., Atrens, D. M. & Shepherd, A. Effects of norepinephrine infused in the paraventricular hypothalamus on energy expenditure in the rat. *Brain Res.* **487**, 79–88 (1989).
20. McGregor, I. S., Menendez, J. A., Atrens, D. M. & Lin, H. Q. Prefrontal cortex α_2 adrenoceptors and energy balance. *Brain Res. Bull.* **26**, 683–691 (1991).
21. Zolovick, A. J., Rossi III, J., Davies, R. F. & Panksepp, J. An improved pharmacological procedure for depletion of noradrenaline: pharmacology and assessment of noradrenaline-associated behaviors. *Eur. J. Pharmacol.* **77**, 265–271 (1982).
22. Himms-Hagen, J. Role of brown adipose tissue thermogenesis in control of thermoregulatory feeding in rats: a new hypothesis that links thermostatic and glucostatic hypotheses for control of food intake. *Proc. Soc. Exp. Biol. Med.* **208**, 159–169 (1995).
23. Erickson, J. C., Clegg, K. E. & Palmiter, R. D. Sensitivity to leptin and susceptibility to seizures of mice lacking neuropeptide Y. *Nature* **381**, 415–418 (1996).
24. Leibowitz, S. F. Brain neuropeptide Y: an integrator of endocrine, metabolic and behavioral processes. *Brain Res. Bull.* **27**, 333–337 (1991).
25. Ahren, B., Mansson, S., Gingerich, R. L. & Havel, P. J. Regulation of plasma leptin in mice: influence of age, high-fat diet, and fasting. *Am. J. Physiol.* (in the press).

Acknowledgements. We thank Sumitomo Pharmaceuticals, Ltd. (Osaka, Japan) for DOPS; Moor Instruments (Millway, UK) for loaning their floLAB laser Doppler perfusion monitor; S. C. Woods and R. J. Seeley for use of their indirect calorimeter; C. H. Warden for the UCP2 clone; P. J. Havel for performing the leptin measurements; B. D. Johnson for help with recording EMGs; A. T. Tafari and C. Kinyoun for technical support; and M. W. Schwartz, D. E. Cummings and J. D. Brunzell for suggestions on the manuscript.

Correspondence and requests for materials should be addressed to R.D.P. (e-mail: palmiter@u.washington.edu).

Hox9 genes and vertebrate limb specification

Martin J. Cohn*†, Ketan Patel*†‡, Robb Krumlauf‡, David G. Wilkinson‡, Jonathan D. W. Clarke* & Cheryll Tickle*

* Department of Anatomy and Developmental Biology, University College London Medical School, Medawar Building, Gower Street, London WC1E 6BT, UK

‡ Division of Developmental Neurobiology, National Institute for Medical Research, The Ridgeway, London NW7 1AA, UK

† These authors contributed equally to this work.

Development of paired appendages at appropriate levels along the primary body axis is a hallmark of the body plan of jawed vertebrates. Hox genes are good candidates for encoding position in lateral plate mesoderm along the body axis^{1,2} and thus for determining where limbs are formed. Local application of fibroblast growth factors (FGFs) to the anterior prospective flank of a chick embryo induces development of an ectopic wing, and FGF applied to posterior flank induces an ectopic leg³. If particular combinations of Hox gene expression determine where wings and legs develop, then formation of additional limbs from flank should involve changes in Hox gene expression that reflect the type of limb induced. Here we show that the same population of flank cells can be induced to form either a wing or a leg, and that induction of these ectopic limbs is accompanied by specific changes in expression of three Hox genes in lateral plate mesoderm. This then reproduces, in the flank, expression patterns found at normal limb levels. Hox gene expression is reprogrammed in lateral plate mesoderm, but is unaffected in paraxial mesoderm. Independent regulation of Hox gene expression in lateral plate mesoderm may have been a key step in the evolution of paired appendages.

In previous experiments, ectopic formation of limbs induced by FGF suggested the presence of two adjacent cell populations in the flank, one with the potential to form a wing and another with the potential to form a leg^{3,4}. To determine directly whether cells that form ectopic wings constitute a separate population from cells that form ectopic legs, a bead soaked in FGF-2 was implanted in the flank either opposite somite (s) 21 (anterior flank), which induces mostly

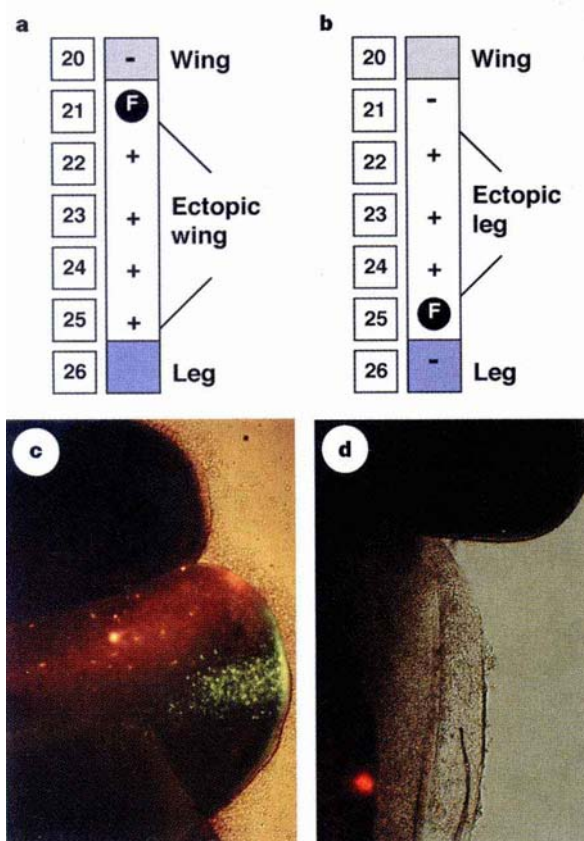


Figure 1 FGF induces bidirectional transformations in flank cell fate to induce ectopic wings and legs. **a, b**, Lateral plate mesoderm with FGF bead (F) placed in anterior flank opposite s21 (**a**), a position that generally results in ectopic wing formation, or in posterior flank opposite s25 (**b**), a position that generally results in ectopic legs. FGF beads were implanted in the lateral plate mesoderm of the flank of stage 14–15 chick embryos and a small patch of cells at one or two positions were labelled with Dil and/or DiA at the time of bead implantation. Embryos were viewed under fluorescence microscopy 48 to 72 h later. (+), Cells labelled at this position contributed to the ectopic limb bud; (–), cells labelled at this position did not contribute. (Anterior is at the top.) **a**, When beads were implanted opposite s21, cells labelled opposite s22 ($n = 1$), 23 ($n = 1$), 24 ($n = 3$), and 25 ($n = 1/2$) were located in the ectopic bud. Cells labelled anterior to the bead, opposite s20 in the normal wing, were not detected in ectopic limbs ($n = 2$). **b**, When beads were placed opposite s25, cells labelled opposite s24 ($n = 1$), 23 ($n = 4$), and 22 ($n = 3/4$) contributed to ectopic limbs. Cells in the anterior flank opposite s21 ($n = 3$) and cells posterior to the bead, opposite s26 ($n = 1$), were not detected in ectopic limbs. **c**, Triple exposure photomicrograph of ectopic limb bud to reveal positions of fluorescent cells 72 h after implantation of a FGF bead opposite s25. At the time of bead implantation, flank cells were labelled with Dil (red) opposite s22, and DiA (green) opposite s23. Cells from both positions have contributed to the ectopic limb bud. Note the extent of cell spread in both anteroposterior and proximodistal axes compared to that in a normal embryo (**d**). **d**, Double-exposure photomicrograph of a normal chick embryo in which flank cells opposite s22 had been labelled with Dil 48 h earlier. Labelled cells remain in the flank as a small patch. There is no contribution to the limb buds and cells remained tightly clustered.

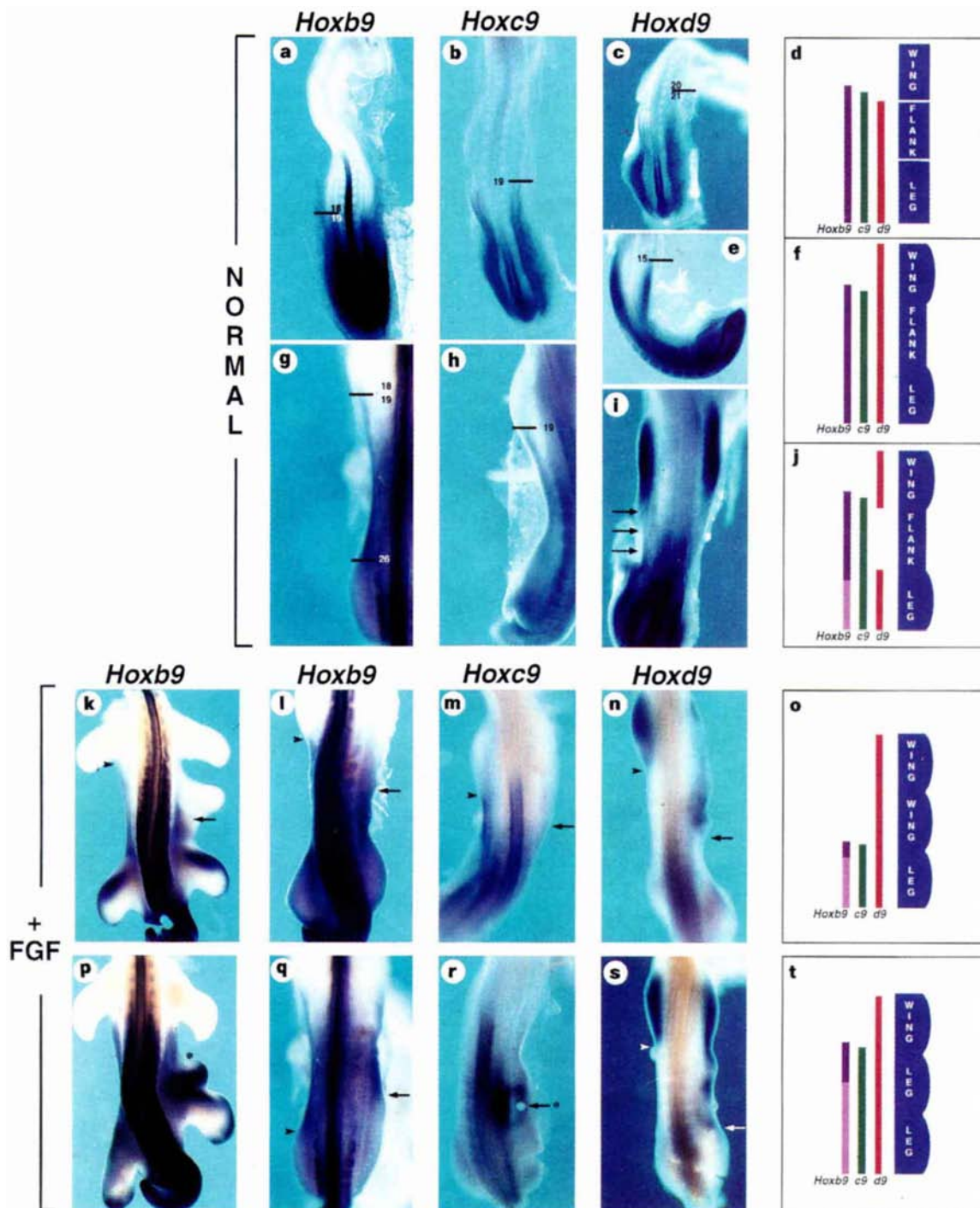
Figure 2 Expression of *Hox* group 9 paralogues in lateral plate mesoderm of normal chick embryos and embryos treated with FGF to induce ectopic limbs. *Hoxb9* (**a, g, k, l, p** and **q**), *Hoxc9* (**b, h, m** and **r**) and *Hoxd9* (**c, e, i, n** and **s**) transcripts detected by whole-mount *in situ* hybridization. Embryos are oriented with anterior at the top and are dorsal views, except for **e**, which is a ventrolateral view. Black lines mark boundaries of expression domains in lateral plate mesoderm of normal embryos; arrows mark boundaries on FGF-treated side (except in **i**) and arrowheads are on the contralateral (untreated side) of embryos. **a–c**, Pre-limb-bud-stage chick embryos (stage 14–15). **a**, Anterior boundary of *Hoxb9* expression in lateral plate mesoderm lies opposite the junction between s18 and 19. **b**, Anterior boundary of *Hoxc9* in lateral plate mesoderm lies opposite s19. **c**, Anterior boundary of *Hoxd9* in lateral plate mesoderm lies opposite s20–21. These patterns are summarized in **d**, in which the staggering of the boundaries in the prospective wing region is shown. **e**, Stage-16 embryo in which limb budding has been initiated and the anterior boundary of *Hoxd9* has shifted anteriorly from s20–21 to s15 at the anterior margin of the emerging wing bud. **f**, Summary of *Hoxb9*, *c9* and *d9* expression at stage 16. **g**, Stage-18, and **h**, stage-19 chick embryos. The anterior boundaries of *Hoxb9* (**g**) and *Hoxc9* (**h**) are still at the same axial level as in **a** and **b**. In **g**, a clear posterior boundary of *Hoxb9* expression has appeared opposite s26 between the strong flank expression and the weaker leg expression, but no such change is seen in *Hoxc9* (**h**), which is uniformly expressed in flank and leg. **i**, Stage 17–18 embryo. *Hoxd9* is now clearly down-regulated in the flank between the wing and leg buds (arrows). **j**, Summary of *Hoxb9*, *Hoxc9* and *Hoxd9* expression at stage 19. Dark and light purple represent intensity of *Hoxb9* expression. **k**, 57 h after application of the FGF bead opposite s23, the posteriorly displaced anterior boundary of *Hoxb9* expression lies at the junction between the ectopic limb bud and the remaining flank (arrow), thereby creating a wing-like pattern in the flank. The boundary of expression on the contralateral side marks the junction between the normal wing bud and anterior flank (arrowhead). **l**, Chick embryo 26 h after implantation of a FGF bead in the

lateral plate mesoderm opposite s23. The anterior boundary of *Hoxb9* expression in the lateral plate mesoderm is shifted posteriorly (arrow) compared to the boundary on the contralateral side (arrowhead) which is unaffected. The posterior expression boundary in the lateral plate mesoderm and expression in the somites are unaffected. **m, n**, Chick embryos 18 h after application of FGF opposite s21. **m**, The anterior boundary of *Hoxc9* has shifted posteriorly over a distance equivalent to 3 somites (arrow). Compare with contralateral side, where anterior expression boundary lies at the posterior edge of the emerging wing bud (arrowhead). **n**, The anterior boundary of *Hoxd9* remains at the anterior limit of the wing bud, but the domain is posteriorly extended into the flank (arrow), whereas on the contralateral side (arrowhead) *Hoxd9* has been switched off. Note that on the treated (right) side, although the anterior boundary is at the appropriate axial level, the expression in the wing is patchy. By this stage, expression in the leg buds has started to downregulate. **o**, Summary of the patterns of *Hoxb9*, *c9* and *d9* expression in the flank when the flank is anteriorized by FGF to take on a wing identity. **p**, 72 h after application of a FGF bead opposite s24, the posterior boundary of *Hoxb9* expression is no longer visible, as expression in the ectopic bud has taken on a leg-like pattern. *Hoxb9* has been downregulated at the anterior margin of the ectopic bud (asterisk), reproducing the pattern of the normal leg but with a reversed anteroposterior pattern. **q**, 17 h after application of a FGF bead opposite s22, the posterior boundary of *Hoxb9* expression in the lateral plate mesoderm on the treated side has shifted anteriorly (arrow). Expression in the somites and the boundary on the contralateral side (arrowhead) are unaffected. **r**, 19 h after application of FGF opposite s25, the boundaries of *Hoxc9* expression are unaffected. *Hoxc9* is upregulated in the flank around the bead (asterisk and arrow). **s**, 18 h after application of FGF opposite s25, *Hoxd9* expression has been extended throughout the entire flank (arrow). **t**, Summary of the patterns of *Hoxb9*, *c9* and *d9* expression in the flank when the flank is posteriorized by FGF to take on a leg identity.

ectopic wings, or opposite s25 (posterior flank), which induces ectopic legs (wing bud develops opposite s15–20, flank opposite s21–25, and leg bud opposite s26–32) (data from ref. 3; Table 1c). A small number of flank cells were then labelled with the fluorescent dyes DiI or DiA at different distances along the anteroposterior axis from the FGF bead, and the positions of labelled cells were examined 48–72 hours later. FGF beads placed in anterior flank induced an ectopic bud posterior to the bead, and flank cells opposite s22 to 25, the entire span of the prospective flank, were incorporated into the bud ($n = 7$; Fig. 1a). Cells just anterior to the FGF bead did not contribute ($n = 2$). FGF beads placed in posterior flank induced an ectopic limb anterior to the bead, and cells opposite s22 to 24 were seen in ectopic buds ($n = 9$), but neither cells opposite s21 ($n = 3$) nor cells immediately posterior to the bead ($n = 1$) were incorpo-

rated (Fig. 1b, c). These results show, rather surprisingly, that when FGF beads are placed at opposite ends of the flank, almost the same population of cells contributes to the ectopic bud, which will form either wing or leg according to the position of the FGF bead. Thus, FGF can induce bidirectional changes in cell fate, and flank cells can be anteriorized to give rise to an ectopic wing or posteriorized to form an ectopic leg.

We next mapped patterns and levels of *Hox* gene expression that characterize different regions of lateral plate mesoderm along the main body axis: prospective wing, prospective leg and the intervening flank. Preliminary observations of *Hoxb9* expression suggested that the anterior boundary of expression in lateral plate mesoderm could be related to limb position, and we therefore focused on expression patterns of *Hox* group 9 paralogues *Hoxb9*,



Hoxc9 and *Hoxd9*. Before initiation of limb budding, anterior expression boundaries of *Hoxb9*, *c9* and *d9* in lateral plate mesoderm are staggered within the prospective wing region, and expression is strong throughout prospective flank and leg regions (Fig. 2a–d). This primary pattern of *Hox* gene expression is established very early in development. *Hoxb9* is first expressed around the posterior primitive streak and the domain then spreads anteriorly until the anterior boundary of expression comes to lie within the prospective wing regions at the four-somite stage (data not shown). When limb budding is initiated, a secondary phase of *Hox* gene expression along the main body axis occurs, in which boundaries of expression undergo realignment and levels of expression change locally. The anterior boundary of *Hoxd9* expression shifts anteriorly from the flank–wing junction to the anterior limit of the wing bud (Fig. 2e). Thus, in the secondary phase, the wing bud expresses *Hoxd9* throughout, and *Hoxb9* and *Hoxc9* posteriorly (Fig. 2f). *Hoxd9* expression in the flank is subsequently downregulated and ultimately switches off (Fig. 2i), so that flank expresses *Hoxb9* and *Hoxc9* but not *Hoxd9* (Fig. 2j). *Hoxb9* transcript levels decrease specifically in the leg bud to produce a posterior boundary that separates strong flank expression and weaker leg expression (Fig. 2g), so that the leg bud expression of *Hoxb9* is low and strong for *Hoxc9* and *Hoxd9* (Fig. 2j). Anterior boundaries of *Hoxb9* and *Hoxc9* remain at the same positions as in the primary phase until the buds are well developed (Fig. 2g, h, j; later stages will be described elsewhere). Thus, these three *Hox* genes are expressed in regionally specific patterns related to limb specification and budding.

To determine whether respecification of flank to form limbs involves changes in *Hox* gene expression, we applied FGF to either anterior flank to induce additional wings or to posterior flank to induce additional legs, and monitored expression of *Hoxb9*, *Hoxc9* and *Hoxd9*. We found specific changes in the pattern of *Hox* gene expression in lateral plate mesoderm, according to the position at which FGF is applied. Changes in expression were almost always confined to boundaries; there were no local patches of downregulation around the FGF bead. FGF beads in anterior flank, which lead to ectopic wings, induced a posterior shift of the anterior boundaries of *Hoxb9* and *Hoxc9* expression (Fig. 2k–m; Table 1a,b). The posterior boundary of *Hoxb9* expression could be seen but was not

affected (Fig. 2l). In contrast to the posterior shift of *Hoxb9* and *c9* anterior boundaries, the anterior boundary of *Hoxd9* was unaffected, but *Hoxd9* expression was maintained in the flank where it would have been switched off during normal development (Fig. 2n). Thus, the combination of *Hox* genes expressed in the flank after anterior FGF application is transformed from the normal flank pattern to a pattern normally found at wing level (Fig. 2o).

FGF beads in posterior flank, which lead to ectopic legs (Table 1c), induced an anterior shift of the posterior boundary of the *Hoxb9* domain (Fig. 2q; Table 1a), stronger *Hoxc9* expression in the flank (Fig. 2r) and *Hoxd9* expression was again maintained in the flank (Fig. 2s); anterior boundaries of *Hoxb9* and *Hoxc9* were unaffected. These changes transformed the normal flank pattern of *Hox* gene expression to a pattern normally found at leg level (Fig. 2t). FGF applied to mid-flank, opposite s22–24, which results in either wing or leg development, induced shifts of both anterior and posterior boundaries of *Hoxb9*, sometimes in the same embryo (Table 1a, b). FGF beads placed at any position within the flank occasionally induced single lateral outgrowths extending from anterior wing to the posterior limit of the flank ($n = 8$), which were associated with posterior shifts in the *Hoxb9* anterior boundary (Table 1b). Changes in the primary pattern of *Hox* gene expression were detected as early as 12 hours after FGF application, when the anterior boundary of *Hoxb9* had shifted posteriorly in the lateral plate mesoderm over a distance of one somite (Table 1a). Activation of the expression of genes that encode signalling molecules controlling outgrowth and patterning occurs much later in ectopic buds. Ectopic *Fgf-8* transcripts are first detectable in flank ectoderm 14–16 hours after FGF application⁵, and ectopic *Shh* is first expressed in flank mesenchyme at 24 hours³. Ectopic limb buds later show specific changes in the complex patterns of *Hox* gene expression characteristic of normal wings (for example, downregulation of *Hoxb9*; Fig. 2k) or legs (for example, asymmetric downregulation of *Hoxb9*; Fig. 2p and Table 1b), indicating that a cascade of gene expression characteristic of either wing or leg has been set in train in flank cells.

Hoxb9, *Hoxc9* and *Hoxd9* have different expression boundaries in neural tube, paraxial mesoderm and lateral plate mesoderm. Boundaries of expression in paraxial mesoderm, like lateral plate

Table 1 Effects of FGF on *Hox* gene expression and morphological patterning in lateral plate mesoderm of the flank

(a) Pattern of *Hoxb9* expression 12–24 h after application of FGF to the flank

Axial level of FGF bead	Number of embryos assayed	Anterior boundary shifted	Posterior boundary shifted	Both boundaries shifted	Boundaries unchanged
Somite 21	6	3	0	0	3*
Somite 22	3	0	1	0	2
Somite 23	6	3	1	0	2*
Somite 24	1	0	0	1	0
Somite 25	3	0	3	0	0

(b) Pattern of *Hoxb9* expression 25–75 h after application of FGF to the flank

Axial level of FGF bead	Number of embryos assayed	Anteriorized wing-like expression pattern	Posteriorized leg-like expression pattern	Expression pattern altered†	Pattern unchanged
Somite 21	9	7 (1‡)	1	0	0
Somite 22	9	3	6	0	0
Somite 23	8	2 (2‡)	1	2	1
Somite 24	9	0 (2‡)	5	1	1
Somite 25	11	2 (3‡)	6	0	0

(c) Morphological pattern of ectopic limb skeleton 10 d after FGF application

Axial level of FGF bead	Number of embryos assayed	Ectopic wings	Ectopic legs	Ectopic limbs§	No ectopic limb
Somite 21	8	4	2	0	2
Somite 22	6	4	2	0	0
Somite 23	17	5	4	6	2
Somite 24	12	3	3	2	4
Somite 25	7	0	5	0	2

* Upregulation within expression domain.

† Expression domain altered but without clear resemblance to wing or leg pattern.

‡ Fin-like outgrowth extending from anterior limit of wing to posterior limit of flank.

§ Ectopic limbs with indeterminate morphology.

mesoderm, are dynamic, whereas in neural tube, anterior expression boundaries of *Hoxb9* and *d9* appear to be fixed early and are generally maintained. FGF application to the flank alters *Hoxb9*, *c9* and *d9* boundaries only in lateral plate mesoderm; expression boundaries in somites and neural tube were unchanged (Fig. 2k–n, p–s). This is consistent with our observation that vertebral identity is not altered in embryos that develop ectopic limbs (data not shown).

Our results are consistent with early primary expression of *Hox* genes in lateral plate mesoderm along the body axis specifying positions where limbs develop. Further evidence for this role for *Hox* genes come from mice lacking *Hoxb5*, in which forelimb position is altered². We found that anterior boundaries of *Hox* group 9 paralogues overlap in the region where the wing will form. Staggered boundaries of *Hox* gene expression are known to be important for specifying positional differences along the body axis, as in the distinct segment types in insects⁶, the chordate neural tube⁷, and the vertebrate axial skeleton⁸. We also found that FGF resets *Hox* expression boundaries in lateral plate mesoderm so that they come to lie in the flank. According to the new pattern of *Hox* gene expression, the same population of flank cells then forms either an additional wing or leg. The fact that an additional limb forms, rather than a shift in position of the nearby normal limb, suggests that limb position has already been specified by a ratchet-like mechanism which irreversibly commits cells. The cascade of gene expression set in train as a result could include the recently identified T-box genes, which are differentially expressed in forelimbs and hindlimbs and are candidates for specifying limb identity⁹. Although positional identity and *Hox* gene expression is reset in the flank so that an additional limb forms, ectopic limbs have reversed polarity and *Shh* is expressed at the anterior^{3,5,10,11}. Thus it appears that polarizing potential in the flank is not reset. Our observation that *Hox* gene expression is altered before activation of *Fgf-8* and *Shh* is reminiscent of the changes in gene expression that enable butterfly prolegs to develop in abdominal segments, when downregulation of *Ubx* and *abd-a* precedes activation of *Antennapedia* (*Antp*) and *Distal-less* (*Dll*) in the prospective limbs¹². Regulation of *Hox* gene expression patterns is unlikely to be a direct effect of FGF because of the timing. Instead, FGF may interfere with systems that regulate positional identity. Examples of genes known to regulate *Hox* gene expression are *trithorax*- and *Polycomb*-group genes (reviewed in ref. 13). Reduced function of *Mixed-lineage leukaemia* (*Mill*) gene, a mouse homologue of *trithorax*, results in failure to maintain *Hox* genes and leads to simultaneous anteriorization and posteriorization of the vertebral column¹⁴. If FGF locally interferes with such regulatory genes in lateral plate mesoderm, this could explain how a single factor induces limbs of different types according to where it is applied. The position of paired appendages has shifted along the body axis during evolution without dramatic reorganization of the axial skeleton^{15–17}, and our finding that *Hox* gene expression can be independently regulated in lateral plate and paraxial mesoderm suggests a mechanism by which this could have occurred. □

Methods

Whole-mount *in situ* hybridization and application of FGF. For whole-mount *in situ* hybridization, embryos were removed from the egg and washed and dissected in 1 × PBS. Embryos were fixed overnight in 4% paraformaldehyde at 4 °C and dehydrated in a series of graded methanol washes. Processing and hybridization were done as described¹⁸, using digoxigenin-labelled riboprobes for the chick genes *Hoxb9*, *Hoxc9* (from C. Tabin) and *Hoxd9* (from D. Duboule). Application of FGF-2 beads was as described³.

Iontophoretic application of DiI and DiA. Small deposits of the lipophilic membrane dye-DiI (Molecular Probes D-282) and DiA (Molecular Probes D-3883) were applied *in ovo* to the lateral plate mesoderm of the flank and prospective limb bud by iontophoresis. Microelectrodes with a tip diameter of ~3 µm were filled at their tips with a small quantity of DiI or DiA (3 mg ml⁻¹ dimethyl formamide) and then backfilled with 1M lithium chloride. These were

then inserted into an electrode holder connected to the positive pole of a 9-volt battery. The electrode was carefully micromanipulated through the ectoderm and into the somatic layer of the lateral plate mesoderm at the appropriate position. The dye was driven out of the electrode by completing the circuit with a second silver wire placed into the egg albumin and attached to the battery's negative terminal. Completing the circuit for ~7 s was sufficient to label a small patch of cells. Dye application was done under a dissecting microscope and the success and position of labelled cells then checked on an epifluorescent microscope fitted with an extra long working distance 20 × objective. Embryos were then allowed to develop for the stated times and examined using Nikon fluorescence microscopy. For subsequent analysis of gene expression in labelled embryos, embryos were fixed and mounted in 4% paraformaldehyde, photographed and dehydrated in graded methanol washes before *in situ* hybridization.

Received 19 December 1996; accepted 3 March 1997.

1. Charite, J., de Graaff, W., Shen, S. & Deschamps, J. Ectopic expression of *Hoxb-8* causes duplication of the ZPA in the forelimb and homeotic transformation of axial structures. *Cell* **78**, 589–601 (1994).
2. Rancourt, D. E., Suzuki, T. & Capechi, M. R. Genetic interaction between *hoxb-5* and *hoxb-6* is revealed by nonallelic noncomplementation. *Genes, Dev.* **9**, 108–122 (1995).
3. Cohn, M. L., Izpisua-Belmonte, J. C., Abud, H., Heath, J. K. & Tickle, C. Fibroblast growth factors induce additional limb development from the flank of chick embryos. *Cell* **80**, 739–746 (1995).
4. Tickle, C. Vertebrate limb development. *Curr. Biol.* **5**, 478–484 (1995).
5. Crossley, P. H., Minowada, G., MacArthur, C. A. & Martin, G. R. Roles for FGF8 in the induction, initiation, and maintenance of chick limb development. *Cell* **84**, 127–136 (1996).
6. Akam, M. *et al.* The evolving role of *Hox* genes in arthropods. *Development* (suppl.) 209–215 (1994).
7. Lumsden, A. & Krumlauf, R. Patterning the vertebrate neuraxis. *Science* **274**, 1109–1115 (1996).
8. Krumlauf, R. *Hox* genes in vertebrate development. *Cell* **78**, 191–201 (1994).
9. Gibson-Brown, J. J. *et al.* Evidence of a role for T-Box genes in the evolution of limb morphogenesis and specification of forelimb-hindlimb identity. *Mech. Dev.* **56**, 93–101 (1996).
10. Ohuchi, H. *et al.* An additional limb can be induced from the flank of the chick embryo by FGF4. *Biochem. Biophys. Res. Commun.* **209**, 809–816 (1995).
11. Vogel, A., Rodriguez, C. & Izpisua-Belmonte, J. C. Involvement of FGF-8 in initiation, outgrowth and patterning of the vertebrate limb. *Development* **122**, 1737–1750 (1996).
12. Warren, R. W., Nagy, L., Selegue, J., Gates, J. & Carroll, S. Evolution of homeotic gene regulation and function in flies and butterflies. *Nature* **372**, 458–461 (1994).
13. Simon, J. Locking in stable states of gene expression: transcriptional control during *Drosophila* development. *Curr. Opin. Cell Biol.* **7**, 376–385 (1995).
14. Yu, B. D., Hess, J. L., Horning, S. E., Brown, G. A. J. & Korsmeyer, S. J. Altered *Hox* expression and segmental identity in *Mill*-mutant mice. *Nature* **378**, 505–508 (1995).
15. Agar, W. E. The development of the anterior mesoderm, and paired fins with their nerves, in *Lepidosiren* and *Protopterus*. *Trans. R. Soc. Edin.* **45**, 611–640 (1907).
16. Goodrich, E. S. *Studies on the Structure and Development of Vertebrates*. (Macmillan, London, 1930).
17. Thorogood, P. & Ferretti, P. *Hox* genes, fin folds and symmetry. *Nature* **364**, 196 (1993).
18. Nieto, M. A., Patel, K. & Wilkinson, D. G. in *Methods in Avian Embryology* 220–235 (ed. Bronner-Fraser, M.) (Academic, San Diego, 1996).

Acknowledgements. This work was supported by the BBSRC (M.C., C.T.) the MRC (K.P., R.K., D.W.), the Wellcome Trust (K.P., J.C.) and the Astor Foundation (C.T.). We thank M. Coates and R. Nittenberg for stimulating discussions, C. Tabin and D. Duboule for *Hox* probes, and L. Wolpert for helpful comments on the manuscript.

Correspondence and requests for materials should be addressed to M.C. (m.cohn@ucl.ac.uk).

Essential role for diacylglycerol in protein transport from the yeast Golgi complex

Brian G. Kearns*, Todd P. McGee*, Peter Mayinger†, Alma Gedvilaite*, Scott E. Phillips*, Satoshi Kagiwada* & Vytas A. Bankaitis*

* Department of Cell Biology, University of Alabama at Birmingham, Birmingham, Alabama 35294-0005, USA

† Zentrum für Molekulare Biologie Heidelberg, Im Neuenheimer Feld 282D-69120, Heidelberg, Germany

Yeast phosphatidylinositol transfer protein (Sec14p) is required for the production of secretory vesicles from the Golgi. This requirement can be relieved by inactivation of the cytosine 5'-diphosphate (CDP)-choline pathway for phosphatidylcholine biosynthesis, indicating that Sec14p is an essential component of a regulatory pathway linking phospholipid metabolism with vesicle trafficking (the Sec14p pathway^{1–6}). Sac1p (refs 7 and 8) is

an integral membrane protein related to inositol-5-phosphatases such as synaptojanin⁹, a protein found in rat brain. Here we show that defects in *Sac1p* also relieve the requirement for *Sec14p* by altering phospholipid metabolism so as to expand the pool of diacylglycerol (DAG) in the Golgi. Moreover, although short-chain DAG improves secretory function in strains with a temperature-sensitive *Sec14p*, expression of diacylglycerol kinase from *Escherichia coli* further impairs it. The essential function of *Sec14p* may therefore be to maintain a sufficient pool of DAG in the Golgi to support the production of secretory vesicles.

The *Sec14p* requirement for yeast Golgi function is bypassed by mutations in seven genes¹⁻⁶. Recessive 'bypass *Sec14p*' mutations define genes encoding negative regulators of the *Sec14p* pathway. Examples include the structural genes for enzymes of the CDP-

choline pathway, *KES1* and *SAC1* (refs 3, 4, 6, 7, 10). *SAC1* encodes an integral membrane protein (*Sac1p*) of the Golgi and endoplasmic reticulum, and *Sac1p* defects result not only in a 'bypass *Sec14p*' phenotype, but also: (1) a cold-sensitivity for growth; (2) a relief of specific yeast actin defects; and (3) a new inositol auxotrophy⁶⁻⁸.

Sac1p shares primary sequence homology with the non-catalytic domains of phosphatidylinositol-4,5-bisphosphate (*PtdInsP₂*)-active inositol-5-phosphatases such as yeast *Inp5p* (J. York and P. Majerus; personal communication) and rat brain synaptojanin⁹. *Sac1p* also shares homology with ORF YNL325c, ORF106c and ORF YOR109w identified by the Yeast Genome Project. Although *Sac1p*-inactivating mutations involve residues conserved among these various homologues (Fig. 1a), deletion of the structural genes for *Inp5p*, ORF YNL325c, ORF106c or ORF YOR109w failed to 'bypass

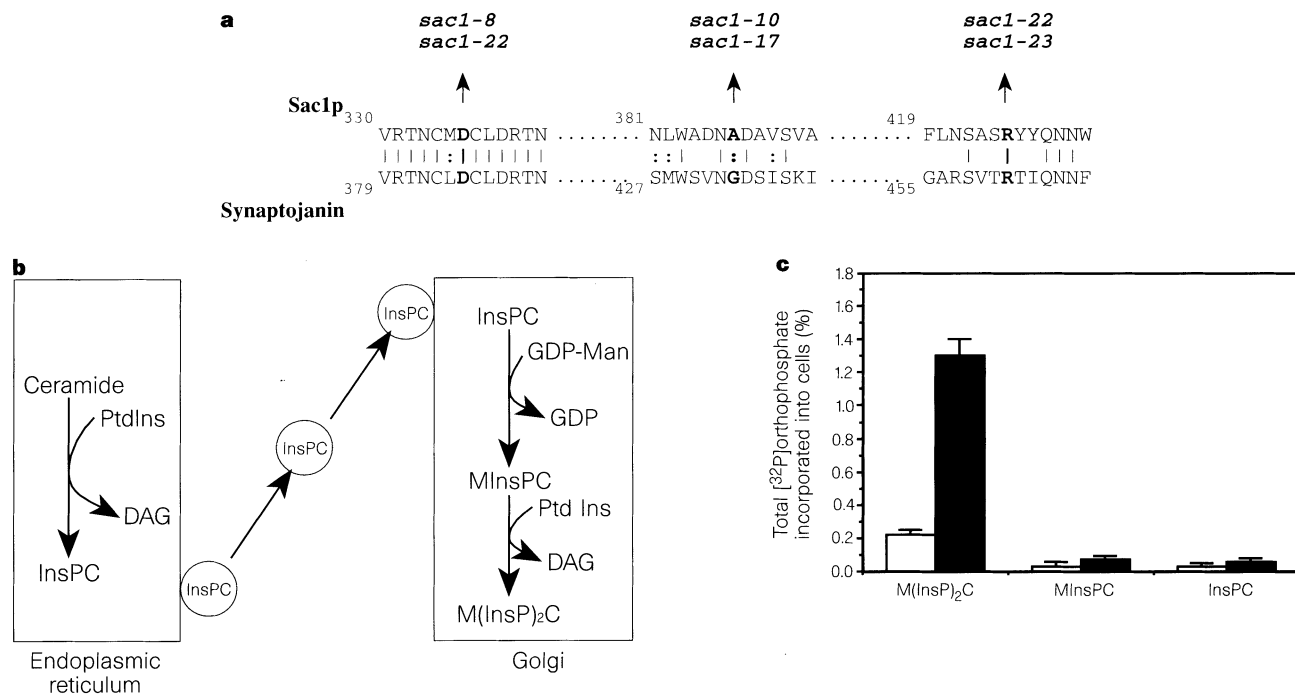


Figure 1 **a** *Sac1p*-inactivating mutations. Designations for *sac1* mutations expressing full-length, but nonfunctional, *Sac1p*^{7,8} are indicated at top with the corresponding amino-acid substitutions. Cognate regions of rat brain synaptojanin are given. Identities (vertical lines) and similarities (double dots) are shown. Mutations represent: D337N, GAT to AAT; A337V, GCA to GTA; and R425H, CGT to CAT. **b**, Yeast inositol-sphingolipid pathway¹¹⁻¹³. Ceramide is decorated with phosphoinositol donated by phosphatidylinositol (*PtdIns*) in the endoplasmic reticulum to generate inositolphosphorylceramides (*InsPC*) and diacylglycerol (*DAG*). *InsPC*s are mobilized through transport vesicles to the Golgi. *InsPC* is mannosylated, with GDP-mannose (*GDP-Man*) as sugar donor, to yield manno-

synylinositolphosphorylceramide (*MInsPC*). Some *MInsPC* is further phosphoinositolyated in the Golgi to yield mannosyldiinositoldiphosphorylceramide (*M(InsP)₂C*); generating one mol Golgi *DAG* per mol *M(InsP)₂C*. **c**, Quantification of inositol sphingolipids as a percentage of total [³²P] incorporated into cells from [³²P]orthophosphate pulse-radiolabelling experiments (*n* > 3) is presented (wild-type, empty bars; *sac1-22*, black bars). The *sac1-22* strain exhibited 70% of wild-type efficiency for [³²P] incorporation (per cell basis), and the marked increase in bulk *M(InsP)₂C* was also readily measured in steady-state radiolabelling experiments (not shown). Strains used: wild-type, CTY182; *sac1-22*, CTY165 (see Methods for detailed genotypes).

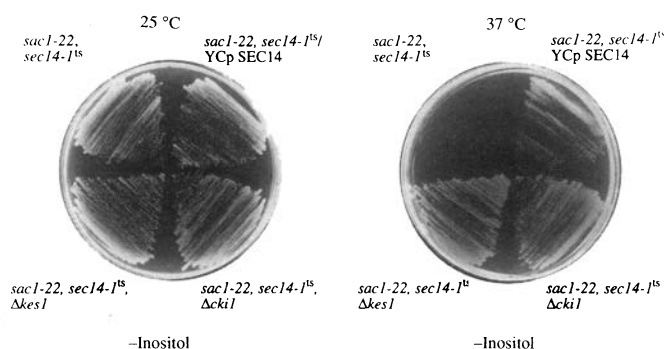


Figure 2 *sac1-22* fails to effect 'bypass *Sec14p*' when inositol is withheld from the medium. Yeast were incubated at the indicated temperatures for 96 h. *sac1-22, sec14-1^{ts}* strains failed to grow in '- Inositol' medium at a restrictive temperature for *sec14-1^{ts}* (37°C), but exhibited inositol prototrophy at a permissive temperature (25°C). Temperature-sensitive inositol auxotrophy of *sac1-22, sec14-1^{ts}* strains was remedied by *SEC14* (for example, YCp(*SEC14*)) or by other 'bypass *Sec14p*' mutations, such as *Δcki1* and *Δkes1*, the structural genes for choline kinase and a yeast homologue of human oxysterol binding protein^{3,10}. Strains used: *sec14-1^{ts}*, *sac1-22*, CTY165; *sec14-1^{ts}, sac1-22/YCpSEC14*, CTY947; *sec14-1^{ts}, sac1-22, Δcki1* CTY952; *sec14-1^{ts}, sac1-22, Δkes1*, CTY953 (see Methods for detailed genotypes).

Sec14p'. Moreover, increased gene dosage of either ORF YNL325c or ORF106c failed to relieve the 'bypass Sec14p' phenotype associated with Sac1p dysfunction (not shown). These data indicate a unique functional interaction between Sac1p and the Sec14p pathway.

[³²P]Orthophosphate radiolabelling experiments demonstrated altered inositol-sphingolipid metabolism in *sac1* strains. Yeast sphingolipids are inositol-phosphoceramide whose headgroups are donated by phosphatidylinositol (PtdIns) in two distinct DAG-producing reactions. The terminal PtdIns-hydrolytic reaction is contained within the Golgi complex¹¹⁻¹³ (Fig. 1b). Remarkably, *sac1* strains experienced a sixfold increase in flux through the inositol-sphingolipid pathway with a commensurate accumulation of mannosyldiinositoldiphosphorylceramide (M(InsP)₂C), the most highly phosphoinositoylated sphingolipid form (Fig. 1c). The altered inositol sphingolipid metabolism experienced by *sac1* strains, that is itself accompanied by accelerated rates of Golgi PtdIns turnover (as shown by the overproduction of M(InsP)₂C; Fig. 1c), provides the first demonstration that Sac1p regulates

inositol phospholipid metabolism^{6,8}. The homology of Sac1p with inositol phospholipid-5-phosphatases is independently consistent with the data.

To examine further the relationship between inositol-sphingolipid metabolism and the 'bypass Sec14p' condition, we took advantage of a particular Sac1p dysfunction caused by the *sac1-22* allele. This Sac1p defect is unique on two counts. First, it does not render yeast auxotrophic for inositol. Second, the accelerated flux through the inositol-sphingolipid pathway in *sac1-22* yeast is inositol-dependent (not shown). Notably, the *sac1-22* allele corrected growth defects resulting from Sec14p deficiency only when yeast were grown in the presence of inositol (Fig. 2). This inositol effect was Golgi-directed as indicated by *sac1-22*-mediated relief of vacuolar proteinase trafficking defects in Sec14p-deficient strains grown in inositol-replete, but not in inositol-free, media (not shown).

Inositol-sphingolipid biosynthesis generates two moles DAG per mole M(InsP)₂C¹¹⁻¹³ (Fig. 1b). [¹⁴C]Acetate pulse-radiolabelling

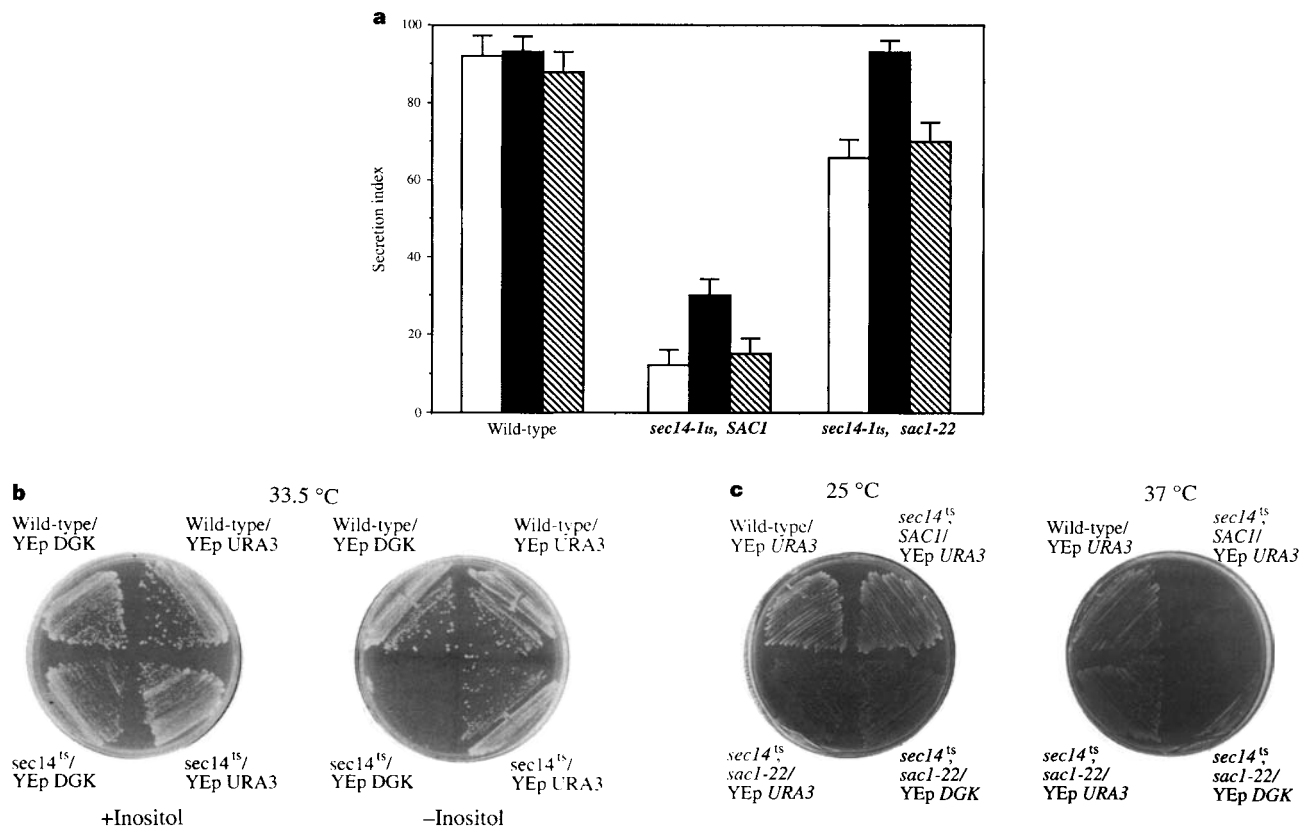


Figure 3 **a**, Invertase secretion as a function of short-chain DAG. Wild-type and *sec14-1^{ts}* strains indicate secretory efficiency (as measured by secretion index at 37 °C²⁵) under Sec14p-proficient and deficient conditions, respectively. *sac1-22*, *sec14-1^{ts}* strains exhibited an intermediate secretion index. These controls are indicated by empty bars. Exogenous di-C₈ DAG (200 μM; black bars) stimulated secretion from *sec14-1^{ts}* and *sac1-22*, *sec14-1^{ts}* strains at 37 °C, but had no effect on wild-type strains (*n* = 5). Stimulation was significant (99% confidence level by Student's *t*-test). Di-C₈ ceramide (200 μM; hatched bars) failed to stimulate secretion. Strains used: wild-type, CTY182; *sec14-1^{ts}*, CTY1-1A; *sec14-1^{ts}*, *sac1-22*, CTY165 (see Methods for detailed genotypes). **b**, DGK exacerbates *sec14-1^{ts}* defects. Yeast were transformed with the insert-less parental plasmid (YEpURA3). Transformants were streaked for isolation on minimal defined medium with inositol (+Inositol; DGK expression repressed) or without inositol (-Inositol; DGK expression derepressed), and incubated at 33.5 °C for 96 h. This temperature is permissive for both *sec14-1^{ts}* and *sec14-1^{ts}*/YEp*PINO1::DGK* strains that do

not express DGK. DGK expression prohibits growth of *sec14-1^{ts}* strains at 33.5 °C, indicating enhancement of *sec14-1^{ts}* defects. Strains used: wild-type/ YEpURA3, CTY948; wild-type/ YEpDGK, CTY949; *sec14-1^{ts}*/ YEpURA3, CTY950; *sec14-1^{ts}*/ YEpDGK, CTY951 (see Methods for detailed genotypes). **c**, DGK compromises suppression of *sec14-1^{ts}* by *sac1-22*. Yeast were transformed with a plasmid driving DGK expression from the powerful and constitutive yeast *PGK* promoter (YEp*Pgk::DGK*), or with the insert-less parental plasmid (YEpURA3). Transformants were incubated on glucose minimal medium lacking uracil at 25 °C or 37 °C for 72 h. DGK expression rendered *sac1-22*, *sec14-1^{ts}* strains incapable of growth at 37 °C, a nonpermissive temperature for *sec14-1^{ts}*. *SEC14* strains expressing DGK grew well at 37 °C. YEpURA3 did not compromise growth of *sac1-22*, *sec14-1^{ts}* strains. Strains used: wild-type/ YEpURA3, CTY948; *sec14-1^{ts}*, *SAC1*/ YEpURA3, CTY950; *sec14-1^{ts}*, *sac1-22*/ YEpURA3, CTY991; *sec14-1^{ts}*, *sac1-22*/ YEpDGK, CTY992 (see Methods for detailed genotypes).

experiments demonstrate that *sac1-22* strains experienced a threefold elevation in bulk cellular DAG content relative to wild-type strains when growth medium contained inositol (the 'bypass Sec14p' condition). This differential in bulk DAG was eliminated when *sac1-22* and wild-type strains were grown in inositol-free medium (not shown); the condition that fails to support *sac1-22*-mediated 'bypass Sec14p' (see above). The relationship between DAG and Sec14p function was further strengthened by the demonstration that di-C8 DAG: (1) effected a 2.8-fold increase in invertase secretion efficiency at 37 °C in yeast strains producing a temperature-sensitive Sec14p (*sec14-1^{ts}*) (Fig. 3a), and (2) enhanced the efficiency with which the *sac1-22* allele improved secretion in Sec14p-deficient strains (Fig. 3a). These effects were specific because short-chain ceramide provided no such enhancement. Finally, di-C8 DAG did not improve invertase secretion in other *sec^{ts}* strains (*sec62-1^{ts}*, *sec18-1^{ts}*, *sec7-1^{ts}*, *sec4-1^{ts}* and *sec9-4^{ts}*).

The reciprocal test of a DAG requirement for yeast Golgi function was to determine whether DAG depletion exacerbated Sec14p defects. Thus, *E. coli* DAG-kinase¹⁴ (DGK) was expressed in *sec14-1^{ts}* strains. As expected, such DGK expression effected elevation of bulk membrane phosphatidic acid (PA) with an accompanying reduction in DAG (not shown). DGK expression clearly reduced the threshold growth temperature for yeast strains producing a temperature-sensitive Sec14p (Fig. 3b), thereby demonstrating that shunting of DAG to PA was detrimental to Golgi function. DGK expression-mediated enhancement of *sec14-1^{ts}* defects was also specific as it had no effect on other *sec^{ts}* strains (*sec62-1^{ts}*, *sec18-1^{ts}*, *sec7-1^{ts}*, *sec4-1^{ts}* and *sec9-4^{ts}*; not shown).

DGK expression was also predicted to compromise the ability of Sac1p deficiency to effect 'bypass Sec14p'. Constitutive expression of DGK from the powerful yeast phosphoglycerate kinase structural gene (*PGK*) promoter strongly diminished the ability of Sac1p-deficiency to alleviate Sec14p defects (Fig. 3c). Thus, contrary to the proposed action of PA as a downstream effector of ADP-ribosylation factor in a mammalian Golgi vesicle assembly pathway¹⁵, DGK-mediated PA production impaired Sec14p-dependent Golgi function. These collective data indicate that reduction of Golgi DAG both: (1) specifically exacerbated Sec14p insufficiencies, and (2) impaired the ability of Sac1p deficiency to effect its usual 'bypass

Sec14p' phenotype.

We suggest that Sec14p functions to preserve an essential Golgi DAG pool (Fig. 4). Sec14p-independent expansion of this DAG pool will result in 'bypass Sec14p'. In constructing models for how Sac1p dysfunction effects 'bypass Sec14p', particularly regarding the involvement of accelerated inositol sphingolipid biogenesis, we wish to address a topological consideration. Mammalian sphingolipids are assembled in the lumen of intracellular organelles¹⁶. If true in yeast, this topology would be seemingly inconsistent with the cytosolic residence of Sec14p; a disposition that limits Sec14p action (and presumably the critical DAG pool) to the cytosolic leaflet of Golgi membranes¹³. However, as DAG flip-flops at appreciable rates between membrane leaflets¹⁷, even an expanded DAG pool generated in the luminal leaflets of Sac1p-deficient Golgi membranes could potentially gain access to the cytosolic leaflet of Sec14p-deficient Golgi and effect 'bypass Sec14p'.

The concept of a Sec14p-dependent DAG pool has several attractive features. First, it identifies a biochemical signal required for progression through the Sec14p pathway. Second, it suggests unique and cooperative functions for both the PtdIns- and phosphatidylcholine (PtdCh)-bound forms of Sec14p in Golgi DAG maintenance (Fig. 4). Third, it accounts for the specific toxicity of the CDP-choline pathway to Golgi function on the basis of its DAG consumption properties³. By focusing on consumption of DAG by the CDP-choline pathway, this model shifts the emphasis from the notion of toxic accumulation of PtdCh; a tenet originally proposed to reconcile the opposing actions of Sec14p and the CDP-choline pathway on Golgi secretory function^{3,4,6}.

In yeast, DAG might represent a precursor for the genuine lipid effector of the Sec14p pathway (which is not PA; Fig. 3b, c). A downstream component of that pathway (for example, Kes1p (ref. 10)) may serve as a Golgi DAG sensor to regulate secretory membranes so as to allow membrane transformation events required for secretory vesicle budding and scission reactions. These concepts emphasize the dynamic interface between phospholipid metabolism and Golgi function.

Finally, Sac1p is also required for efficient ATP transport into the endoplasmic reticulum lumen¹⁸. This raises the issue of how loss of

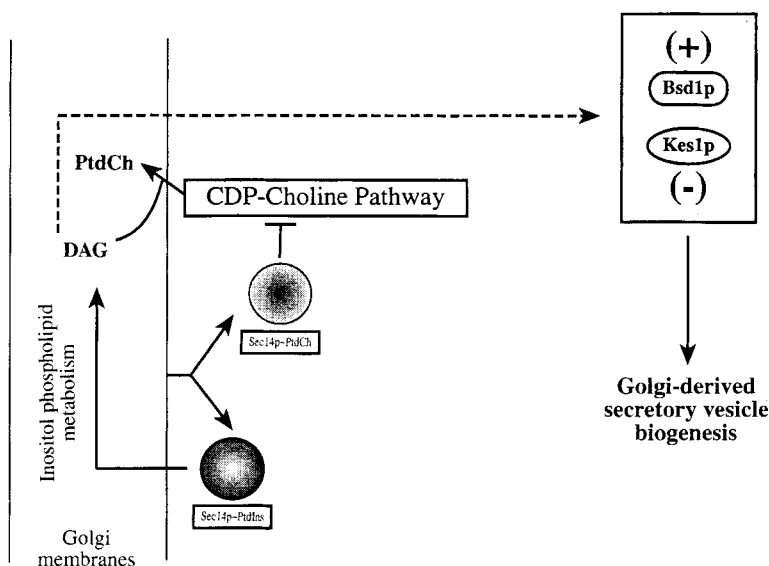


Figure 4 Sec14p preserves Golgi DAG. PtdCh-bound Sec14p (Sec14p~PtdCh) negatively regulates Golgi CDP-choline pathway activity and prevents DAG consumption by this PtdCh biosynthetic pathway^{4,5}. PtdIns-bound Sec14p (Sec14p~PtdIns) stimulates PtdIns metabolism through which Golgi DAG may be replenished. This latter action of Sec14p~PtdIns is consistent with the requirement for PtdIns-transfer activity in mammalian PtdInsPs-mediated

rescue of *sec14-1^{ts}* defects²⁹, and with accelerated PtdIns-turnover representing a mechanism for 'bypass Sec14p'³⁰. Downstream regulators of the Sec14p pathway, the *BSD1* and *KES1* gene products (positive (+) and negative (-) effectors in the Sec14p pathway, respectively¹⁰) represent candidate DAG-responsive activities.

such an activity might relate to 'bypass Sec14p'. We found that endoplasmic reticulum microsomes prepared from *sac1-22* yeast strains exhibited wild-type rates of ATP import, and that overexpression of this defective Sac1p gene product resulted in a proportional overproduction of ATP import activity. Yet, overexpression of the *sac1-22* gene product did not diminish the 'bypass Sec14p' phenotype normally associated with it (not shown). The uncoupling of 'bypass Sec14p' from *sac1*-associated ATP import defects indicates that the relationship between Sac1p and inositol phospholipid metabolism is directly relevant to Sec14p pathway activity whereas the ATP import function of Sac1p is not. □

Methods

Yeast strains. Wild-type strain CTY182 (*Mata ura3-52, Δhis3-200, lys2-801, SEC14, SAC1*); *Δsac1* strain CTY244 (CTY182 *Δsac1*); *sec14-1^{ts}* strain CTY1-1A (*Mata ura3-52 Δhis3-200, lys2-801, sec14-1^{ts}*); *sec14-1^{ts}, sac1-22* strain CTY165 (*Mata ura3-52, Δhis3-200, ade2-101, sec14-1^{ts}, sac1-22*); *sec14-1^{ts}, sac1-22/YCpSEC14* strain CTY947 (CTY165/YCpSEC14); *sec14-1^{ts}, sac1-22, Δcki* strain CTY952 (CTY165 *Δcki*); *sec14-1^{ts}, sac1-22, Δkes1* strain CTY953 (CTY165 *Δkes1*); *SEC14/YEpURA3* strain CTY948 (CTY182/YEpURA3); *SEC14/YEpPINO1::DGK* strain CTY949 (CTY182/YEpPINO1::DGK); *sec14-1^{ts}/YEpURA3* strain CTY950 (CTY1-1A/YEpURA3); *sec14-1^{ts}/YEpPINO1::DGK* strain CTY951 (CTY1-1A/YEpPINO1::DGK); *SEC14/YEpPpGK::DGK* strain CTY949 (CTY182/YEpPpGK::DGK); *sac1-22, sec14-1^{ts}/YEpURA3* strain CTY991 (CTY165/YEpURA3); *sac1-22, sec14-1^{ts}/YEpPpGK::DGK* strain CTY992 (CTY165/YEpPpGK::DGK).

Media and genetic techniques. The minimal defined media and complex media (YPD; supplemented with glucose to a final concentration of 2%) used in all of the experiments in this work have been described⁴. Yeast transformation¹⁹ and gene disruption²⁰ methods are standard. The abilities of individual or combinatorial gene disruption mutations to effect 'bypass Sec14p' were assessed by plasmid shuffle as described in detail elsewhere²¹. Sac1p-inactivating mutations were recovered from *sac1* strains by gap repair²². **Lipid analyses and quantification.** In experiments where phospholipid species were identified and quantified, cells were grown to mid-logarithmic phase in inositol-containing minimal medium and radiolabelled with [³²P]orthophosphate for 20 min at 25°C⁴. Lipids were extracted and resolved by two-dimensional paper chromatography using previously described solvents^{11,12}. Extracts derived from equal numbers of cells were loaded on each chromatogram, and specific [³²P]-radiolabelled phospholipids were identified, excised, and quantified by scintillation counting⁴.

Radiolabelled acetate was used in experiments where DAG and PA were monitored. To this end, cells were grown to mid-logarithmic growth phase (absorbance at 600 nm = 0.6–1.0) in minimal medium with or without inositol and radiolabelled with [2-¹⁴C]-acetate (2 μCi ml⁻¹) for 20 min at 25°C with shaking⁴. DAG was resolved by thin-layer chromatography as previously described^{23,24}. [¹⁴C]-Radiolabelled lipids were quantified by phosphorimaging⁴.

Invertase secretion index. Yeast strains were grown in YPD medium⁴ at 25°C and incubated with 200 μM 1,2-dioctanoyl-*sn*-glycerol (diC₈-DAG) or di-C₆-ceramide in dimethylsulphoxide (DMSO) for 15 min before simultaneous shift to 37°C (to impose Sec14p deficiency) and low glucose (0.1%) YPD (to induce synthesis of secretory invertase). Mock controls were treated with the same amount of DMSO as diC₈-DAG-challenged cultures. Invertase secretion indices were determined as described²⁵.

Construction of P_{INO1}::DGK and P_{pGK}::DGK. Polymerase chain reaction (PCR) technology (using the appropriate primers and *E. coli* genomic DNA as template) was used to place expression of the *E. coli* DGK gene under yeast promoter control. The primers were designed to flank the entire DGK open reading frame, including the cognate initiating ATG and terminating TAA codons, and were clamped with *HindIII* and *SphI* sites at the 5' ends of the forward and reverse primers, respectively. The resulting PCR fragment (representing DGK nucleotide sequence from -1 to +337; ATG initiation codon defining base +1) was digested with *HindIII* and *SphI* and ligated into the corresponding half-sites of pRE247 (ref. 26) to generate plasmid pRE626. The *INO1* and *PGK* promoters were PCR-amplified from total yeast genomic DNA such that the amplified promoters were flanked with an upstream *EcoRI* site

and a downstream *HindIII* site, respectively. The promoter fragments (*INO1* bases -330 to -1; *PGK* bases -840 to -1) were restricted with *EcoRI* and *HindIII* and individually ligated into the cognate half-sites of pRE626. The corresponding DGK expression cassettes were then excised as *EcoRI*-*SphI* restriction fragments and ligated singly into the multi-copy yeast shuttle vector YEplac195 (ref. 27) generating pCTY84 (*P_{INO1}::DGK*) and pCTY85 (*P_{pGK}::DGK*). Translational termination of the expressed DGK polypeptide is, in both cases, mediated by the native DGK termination codon.

Received 15 January; accepted 26 February 1997.

- Bankaitis, V. A., Malehorn, D. E., Emr, S. D. & Greene, R. The *Saccharomyces cerevisiae* SEC14 gene encodes a cytosolic factor that is required for transport of secretory proteins from the yeast Golgi complex. *J. Cell Biol.* **108**, 1271–1281 (1989).
- Bankaitis, V. A., Aitken, J. R., Cleves, A. E. & Dowhan, W. An essential role for a phospholipid transfer in yeast Golgi function. *Nature* **347**, 561–562 (1990).
- Cleves, A. E. *et al.* Mutations in the CDP-choline pathway for phospholipid biosynthesis bypass the requirement for an essential phospholipid transfer protein. *Cell* **64**, 789–800 (1991).
- McGee, T. P., Skinner, H. B., Whitters, E. A., Henry, S. A. & Bankaitis, V. A. A phosphatidylinositol transfer protein controls the phosphatidylcholine content of yeast Golgi membranes. *J. Cell Biol.* **124**, 273–287 (1994).
- Skinner, H. B. *et al.* Phosphatidylinositol transfer protein stimulates yeast Golgi secretory function by inhibiting choline-phosphate cytidylyltransferase activity. *Proc. Natl Acad. Sci. USA* **92**, 112–116 (1995).
- Cleves, A. E., McGee, T. P. & Bankaitis, V. A. Phospholipid transfer proteins: a biological debut. *Trends Cell Biol.* **1**, 30–34 (1991).
- Cleves, A. E., Novick, T. P. & Bankaitis, V. A. Mutations in the SAC1 gene suppress defects in yeast Golgi and yeast actin function. *J. Cell Biol.* **109**, 2939–2950 (1989).
- Whitters, E. A., Cleves, A. E., McGee, T. P., Skinner, H. B. & Bankaitis, V. A. SAC1p is an integral membrane protein that influences the cellular requirement for phospholipid transfer protein function and inositol in yeast. *J. Cell Biol.* **122**, 79–94 (1992).
- McPherson, P. S. *et al.* A presynaptic inositol-5-phosphatase. *Nature* **379**, 353–357 (1996).
- Fang, M. *et al.* Ksp1 shares homology with human oxysterol binding protein and participates in a novel regulatory pathway for yeast Golgi-derived transport vesicle biogenesis. *EMBO J.* **15**, 6447–6459 (1996).
- Becker, G. W. & Lester, R. L. Biosynthesis of phosphoinositol-containing sphingolipids from phosphatidylinositol by a membrane preparation from *Saccharomyces cerevisiae*. *J. Bacteriol.* **142**, 747–754 (1980).
- Puoti, A., Despons, C. & Conzelmann, A. Biosynthesis of mannosylinositolphosphoceramides in *Saccharomyces cerevisiae* is dependent on genes controlling the flow of secretory vesicles from the endoplasmic reticulum to the Golgi. *J. Cell Biol.* **113**, 515–525 (1991).
- Lester, R. L. & Dickson, R. C. Sphingolipids and inositol containing headgroups. *Adv. Lipid Res.* **26**, 253–271 (1993).
- Lightner, V. A., Bell, R. M. & Modrich, P. The DNA sequences encoding *plsB* and *dgk* loci of *Escherichia coli*. *J. Biol. Chem.* **258**, 10856–10861 (1983).
- Kistakis, N. T., Brown, H. A., Waters, M. G., Sternweis, P. C. & Roth, M. G. Evidence that phospholipase D mediates ADP ribosylation factor-dependent formation of Golgi coated vesicles. *J. Cell Biol.* **134**, 295–306 (1996).
- Lipsky, N. G. & Pagano, R. E. Intracellular translocation of fluorescent sphingolipids in cultured fibroblasts: endogenously synthesized sphingomyelin and glucocerebroside analogues pass through the Golgi apparatus en route to the plasma membrane. *J. Cell Biol.* **100**, 27–34 (1985).
- Ganong, B. R. & Bell, R. M. Transmembrane movement of phosphatidylglycerol and diacylglycerol sulphydryl analogues. *Biochemistry* **23**, 4977–4983 (1984).
- Mayinger, P., Bankaitis, V. A. & Meyer, D. I. Sac1p mediates the adenosine triphosphate transport into yeast endoplasmic reticulum that is required for protein translocation. *J. Cell Biol.* **131**, 1377–1386 (1995).
- Ito, H., Fukuda, Y., Murata, K. & Kimura, A. *J. Bacteriol.* **153**, 163–168 (1983).
- Rothstein, R. J. One step gene disruption in yeast. *Methods Enzymol.* **101**, 202–211 (1983).
- Lopez, M. C. *et al.* A phosphatidylinositol/phosphatidylcholine transfer protein is required for differentiation of the dimorphic yeast *Yarrowia lipolytica* from the yeast to the mycelial form. *J. Cell Biol.* **124**, 113–127 (1994).
- Orr-Weaver, T. L., Szostak, J. L. & Rothstein, R. J. Genetic applications of yeast transformation with linear and gapped plasmids. *Methods Enzymol.* **101**, 202–211 (1983).
- Buttke, T. M. & Pyle, A. L. Effects of unsaturated fatty acid deprivation on neutral lipid synthesis in *Saccharomyces cerevisiae*. *J. Bacteriol.* **152**, 747–756 (1982).
- Steiner, M. R. & Lester, R. L. *In vitro* studies of phospholipid biosynthesis in *Saccharomyces cerevisiae*. *Biochim. Biophys. Acta* **260**, 222–243 (1972).
- Salama, S. R., Cleves, A. E., Malehorn, D. E., Whitters, E. A. & Bankaitis, V. A. Cloning and characterization of *Kluyveromyces fragilis* SEC14, a gene whose product stimulates Golgi secretory function in *Saccharomyces cerevisiae*. *J. Bacteriol.* **172**, 4510–4521 (1990).
- Skinner, H. B., Alb, J. G., Jr, Whitters, E. A., Helmkamp, G. M. Jr & Bankaitis, V. A. Phospholipid transfer activity is relevant to but not sufficient for the essential function of the yeast SEC14 gene product. *EMBO J.* **12**, 4775–4784 (1993).
- Gietz, R. D. & Sugino, A. New yeast-*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six base-pair restriction sites. *Gene* **74**, 527–534 (1988).
- Lopes, J. M. & Henry, S. A. Interaction of *trans* and *cis* regulatory elements in the *INO1* promoter of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **14**, 3987–3994.
- Alb, J. G., Jr, Gedvilaite, A., Cartee, R. T., Skinner, H. B. & Bankaitis, V. A. Mutant rat phosphatidylinositol/phosphatidylcholine transfer proteins specifically defective in phosphatidylinositol transfer: implications for the regulation of phospholipid transfer activity. *Proc. Natl Acad. Sci. USA* **92**, 8826–8830 (1995).
- Kagiwada, S. *et al.* The yeast *BSD2-1* mutation influences both the requirement for phosphatidylinositol transfer protein function and derepression of phospholipid biosynthetic gene expression in yeast. *Genetics* **143**, 685–697 (1996).

Acknowledgements. This work was supported by grants from the NIH and the American Cancer Society to V.A.B., B.G.K., A.G. and S.K. were supported by an NSF predoctoral Neurobiology training grant, a grant from the Council for Tobacco Research, and a postdoctoral fellowship from the Human Frontier Science Program, respectively. P.M. was supported by a grant from the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to V.A.B. (e-mail: VABANKAITIS@bmg.bhs.uab.edu).

Prediction of bone density from vitamin D receptor alleles

**Nigel A. Morrison, Jian Cheng Qi, Akifumi Tokita,
Paul J. Kelly, Linda Crofts, Tuan V. Nguyen,
Philip N. Sambrook & John A. Eisman**

Nature 367, 284–287 (1994)

In order to expand our original observations of a relationship between bone density and vitamin D receptor (VDR) genotype and to examine other genes potentially involved in bone biology, we set out to recruit a larger set of identical and non-identical twins. On re-analysis of these newly collected and re-genotyped samples, we found reduced correlation of the VDR genotype with bone density in this larger sample of twin pairs. We made a formal announcement of this finding at the World Congress on Osteoporosis in Amsterdam on 20 May 1996.

As this larger twin sample included some of the twins reported earlier, we re-examined the original samples and found that in a proportion of these twins the genotype (by PCR) on new leukocyte DNA samples differed from those obtained on the earlier leukocyte DNA samples (also by PCR). It seems most likely that the misclassifications arose from mis-genotyping of DNA samples between extraction and PCR analysis. We emphasize that the other major

part of the paper, showing a genotype effect in a population sample, is not affected by such mis-genotyping.

A role of the VDR alleles in bone biology has been reproduced in a wide range of clinical and physiological studies. Other studies have found an effect but in the reverse direction and some have found no effect on bone density or turnover or on osteoporosis prevalence. Thus, while there is disagreement about the strength of the effect, these and other studies in several population samples support the role of the VDR in the polygenic inheritance of bone density. □

Diverse sources of hippocampal unitary inhibitory postsynaptic potentials and the number of synaptic release sites

Eberhard H. Buhl, Katalin Halasy & Peter Somogyi

Nature 368, 823–828 (1994)

The scale bar for all panels of Fig. 4 of this Article was 0.4 μm , and not 0.2 μm as published. □

KNOW YOUR COPY RIGHTS RESPECT OURS

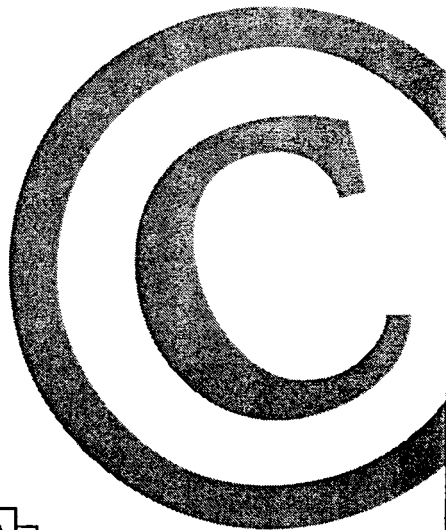
The publication you are reading is protected by copyright law. This means that the publisher could take you and your employer to court and claim heavy legal damages if you make unauthorised photocopies from these pages.

Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

The Copyright Licensing Agency (CLA) is an organisation which issues licences to bring photocopying within the law. It has designed licensing services to cover all kinds of special needs in business, education, and government.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931
Fax: 0171 436 3986